

Faculdade de Engenharia da Universidade do Porto



Previsão de carga – Comparação de técnicas

Tiago Miguel Bastos Barros

Dissertação realizada no âmbito do
Mestrado Integrado em Engenharia Eletrotécnica e de Computadores
Major Energia

Orientador: Professor Doutor José Nuno Fidalgo

Julho de 2014

A Dissertação intitulada

“Previsão de Carga - Comparação de técnicas”

foi aprovada em provas realizadas em 16-07-2014

o júri



Presidente **Professora Doutora Maria Helena Osório Pestana de Vasconcelos**
Professora Auxiliar do Departamento de Engenharia Eletrotécnica e de
Computadores da Faculdade de Engenharia da Universidade do Porto



Professor Doutor Filipe Miguel Tavares Azevedo
Professor Adjunto do Departamento de Engenharia Electrotécnica do Instituto
Superior de Engenharia do Porto



Professor Doutor José Nuno Moura Marques Fidalgo
Professor Associado do Departamento de Engenharia Eletrotécnica e de
Computadores da Faculdade de Engenharia da Universidade do Porto

O autor declara que a presente dissertação (ou relatório de projeto) é da sua exclusiva autoria e foi escrita sem qualquer apoio externo não explicitamente autorizado. Os resultados, ideias, parágrafos, ou outros extratos tomados de ou inspirados em trabalhos de outros autores, e demais referências bibliográficas usadas, são corretamente citados.



Autor - Tiago Miguel Bastos Barros

Faculdade de Engenharia da Universidade do Porto

Resumo

A previsão de carga é um processo indispensável tanto para planeamento como operação dos sistemas de energia elétrica. Além disso, representa também um fator de vantagem na gestão de ativos das empresas que participam no mercado de energia, permitindo uma gestão mais eficiente e económica do fornecimento de energia elétrica.

Conseguir antever eficazmente e com o menor erro possível o comportamento da carga, possibilita um planeamento e operação do sistema elétrico mais eficiente e lucrativo. Daí a pertinência em estudar modelos/técnicas de previsão de carga e avaliar o seu desempenho. Em condições reais de operação, o tempo de resposta é tão crítico quanto a sua precisão, pelo que é importante destacar o esforço computacional necessário em estudos de sistemas de energia complexos, onde a redução dos tempos de processamento não é apenas uma questão de eficiência computacional, mas fundamentalmente de precisão/eficácia de resposta.

A partir de uma boa previsão, é possível decidir com segurança pelo aumento de geração de energia numa determinada hora do dia, ou até determinar quantas centrais elétricas necessitam ser construídas num horizonte de longo prazo, tendo em conta a expansão do sistema e o aumento da procura de energia. Contudo, são diversas as variáveis que afetam o consumo de energia num sistema elétrico, o que torna a previsão bastante complexa. A temperatura, a hora do dia, os dias da semana, as estações do ano e até alguns fatores socioeconómicos são exemplos de fatores que condicionam o consumo.

Vários são os modelos capazes de fazer uma previsão de carga, com maior ou menor precisão. No âmbito desta dissertação, pretende-se comparar o desempenho de diferentes técnicas de previsão de carga de curto prazo, nomeadamente:

I) Análise clássica (Box-Jenkins) de séries temporais;

II) Redes neuronais;

III) Máquinas de suporte vetorial (*SVM – Support Vector Machines*, na literatura anglo-saxónica).

A análise clássica (ponto I) permitirá estabelecer um patamar de referência em termos de desempenho, que se espera ultrapassar com as duas técnicas seguintes. Na comparação entre técnicas, serão confrontados diferentes tipos de índices de desempenho (erros), bem como tempos de treino, flexibilidade e capacidade de filtragem de dados anómalos.

Palavras-chave:

Previsão de carga, séries temporais Box-Jenkins, redes neuronais, SVM.

Abstract

Load forecasting is an indispensable process for both planning and operation of electric power systems. In addition, is also an advantage factor in asset management companies that participate in the energy market, enabling a more efficient and economic supply of electricity.

Effectively being able to predict the behavior of load demand and achieve the lowest possible error, enables a more efficient and profitable planning and operation of the electrical system, thus the relevance of studying models/techniques for load forecasting and evaluating their performance. In real operating conditions, the response time is so critical as to its accuracy, therefore it is important to highlight the computational effort required in studies of complex energy systems where the reduction in processing times is not just a matter of computational efficiency, but fundamentally of the response accuracy/effectiveness.

From a good prediction, it is possible to decide with safety to increase power generation at a certain time of day, or even determine how many power plants need to be built on a longer time horizon, taking into account the expansion of the power system and the increasing of energy demand. However, there are several variables that affect energy consumption in a power system, like the temperature, hour of the day, day of the week, the season and even some social economical factors. These exemplify some of the issues that affect the power system and load forecasting.

Several models are able to make a load forecasting with higher or lower error rates. Within this dissertation, the intention is to compare the performance of different techniques of short-term load forecasting, including:

- I) classical analysis (Box- Jenkins) time series;
- II) neural networks;
- III) support vector machines - SVM (in Anglo-Saxon literature).

Classical analysis (Section I) will establish a level of reference in terms of performance, which is expected to overcome with the following two techniques. In the confrontation of the different techniques, several types of performance indices (errors), as well as practice times, flexibility and data filtering outliers will be compared.

Keywords:

Load forecasting, Box-Jenkins time series, neural networks, SVM.

Agradecimentos

No desenvolvimento desta dissertação, contribuiu o apoio, incentivo e conselhos úteis que o meu orientador de tese proporcionou, o Professor Doutor José Nuno Fidalgo.

Quanto aos agradecimentos, não sendo muitos, são sentidos, nomeadamente para a minha Mãe que sempre me acompanhou e acarinhou, para a Tia Rosário, Tia Alzira, Avós, primas(os) e amigos: Hugo, Ganso, José Luís, João Medeiros, Micas, Nuno, Rémy, Meirinhos, Tomás, Vidal e Ivo, por todo o incentivo, apoio, amizade e convívio proporcionado.

Um agradecimento especial para a minha namorada Joana, por se dedicar a mim e por me ajudar a ultrapassar os dias mais difíceis, fazendo-me sorrir mesmo quando não apeteceria e por toda a confiança e carinho que tem diariamente comigo.

Índice

Resumo	iii
Abstract.....	v
Agradecimentos	vii
Índice	ix
Lista de figuras	xii
Lista de tabelas.....	xv
Abreviaturas e símbolos	xvii
Capítulo 1	1
Introdução.....	1
1.1 Considerações Gerais	1
1.2 Enquadramento	3
1.3 Objetivos.....	5
1.4 Estrutura da Dissertação	5
1.5 Software Utilizado	6
Capítulo 2	8
Métodos de Previsão.....	8
2.1 Modelo Clássico (Box-Jenkins)	8
2.2 Redes Neurais	11
2.3 Máquinas de Suporte Vetorial (SVM).....	14
2.4 Avaliação do desempenho dos modelos	17
2.5 Conclusão.....	18
Capítulo 3	20
Previsão baseada em Box-Jenkins	20
3.1 Verificação da condição de estacionariedade	20
3.2 Identificação do modelo	22
3.3 Estimação de parâmetros	24
3.4 Comparação do consumo real com o estimado	25
3.5 Consideração da sazonalidade	27
3.6 Avaliação do desempenho do modelo	29
3.7 Conclusão.....	30
Capítulo 4	32

Previsão com Redes Neuronais	32
4.1 <i>Software</i> utilizado	32
4.2 Aplicação de RN em estudos de previsão	34
4.3 Estandarização dos dados de entrada	37
4.4 Caracterização do consumo face ao dia de semana e hora	40
4.4.1 Consideração dos feriados	41
4.5 Avaliação do desempenho do modelo	42
4.6 Conclusão.....	44
Capítulo 5	46
Previsão com SVM.....	46
5.1 <i>Software</i> utilizado	46
5.2 Resultados	47
5.3 Conclusão.....	49
Capítulo 6	51
Comparação das Técnicas.....	51
Capítulo 7	54
Conclusão.....	54
Trabalhos futuros	55
Referências.....	56

Lista de figuras

Figura 1.1 – Diagrama de carga do dia 24 de novembro de 2012 [1]	3
Figura 1.2 – Diagrama de carga do dia 9 de outubro de 2013 [1]	4
Figura 1.3 – Fontes de produção de energia em Portugal [2]	4
Figura 2.1 – Fluxograma para a construção de um modelo ARIMA	11
Figura 2.2 – Constituição de um neurónio biológico [6]	11
Figura 2.3 – Esquema exemplificativo de uma rede neuronal	12
Figura 2.4 – Modelo genérico de um neurónio artificial.....	12
Figura 2.5 – Tipos de funções de ativação típicas em RN	13
Figura 2.6 – Margem de separação p para o caso de duas classes linearmente separáveis.....	15
Figura 2.7 – Gráfico da função de penalização para um determinado ϵ	16
Figura 3.1 – Série relativa ao consumo do ano de 2005	20
Figura 3.2 – Série de consumo, com a média e o desvio padrão em janelas de 30 dias	21
Figura 3.3 – Série de consumo estacionária.....	22
Figura 3.4 – Gráfico ACF da série estacionária	23
Figura 3.5 – Gráfico PACF da série estacionária	23
Figura 3.6 – Correlograma do Consumo Real Vs Estimado	25
Figura 3.7 – Comparação do Consumo Real com o Estimado	25
Figura 3.8 – Gráfico ACF do resíduo.....	26
Figura 3.9 – Gráfico PACF do resíduo.....	26
Figura 3.10 – Comparação do Consumo Real com o Estimado	28
Figura 4.1 – Esquema representativo da arquitetura da RN	32
Figura 4.2 – Percentagens a definir para o treino, validação e teste da RN	33
Figura 4.3 – Parâmetros de progresso durante o treino da RN.....	33
Figura 4.4 – Esquema da RN definida para o treino	35
Figura 4.5 – Resultados para o coeficiente de correlação “R” do treino, validação e teste da RN	35
Figura 4.6 – Parâmetros de progresso durante o treino da RN.....	36
Figura 4.7 – Comparação do Consumo Real com a Rede Neuronal	36
Figura 4.8 – Resultados para o coeficiente de correlação “R” do treino, validação e teste da RN estandardizada	38
Figura 4.9 – Parâmetros de progresso durante o treino da RN estandardizada	38
Figura 4.10 – Comparação do Consumo Real com a Rede Neuronal estandardizada.....	39

Figura 4.11 – Comparação do consumo real e previsto no feriado de 27 de março 2005	42
Figura 5.1 – Comparação entre o consumo real e o previsto	48

Lista de tabelas

Tabela 3.1 – Comparação dos vários modelos de previsão	29
Tabela 4.1 – Variáveis a considerar no treino da RN (1º modelo).....	34
Tabela 4.2 – Variáveis a considerar no treino da RN (2º modelo).....	39
Tabela 4.3 – Comparação de desempenho	43
Tabela 5.1 – Comparação dos diferentes modelos SVM	48
Tabela 5.2 – Comparação de desempenho	49
Tabela 6.1 – Comparação de desempenho dos melhores modelos de cada método	51

Abreviaturas e símbolos

Lista de abreviaturas

ACF	<i>Auto Correlation Factor</i>
AR	Auto-regressivo
ARIMA	<i>Auto Regressive Integrated Moving Average</i>
BJ	Box-Jenkins
DS	Dia de semana
FEUP	Faculdade de Engenharia da Universidade do Porto
MA	<i>Moving Average</i>
MAPE	Erro médio percentual absoluto
PACF	<i>Partial Auto Correlation Factor</i>
PRE	Produção em Regime Especial
R	Coefficiente de correlação
REN	Rede Elétrica Nacional
RMSE	Erro médio quadrático
SARIMA	<i>Seasonal Auto Regressive Integrated Moving Average</i>
SEE	Sistema Elétrico de Energia
SVM	<i>Support Vector Machines</i>

Lista de símbolos

W	Watt
MW	Megawatt
MWh	Megawatt-hora

Capítulo 1

Introdução

No âmbito do Mestrado Integrado em Engenharia Eletrotécnica e Computadores, da Faculdade de Engenharia da Universidade do Porto, foi desenvolvida esta dissertação com o objetivo de analisar e comparar diferentes técnicas quando aplicadas a previsão de carga de curto prazo.

Neste capítulo inicial, apresenta-se uma abordagem geral à estrutura do sistema elétrico nacional e à importância e pertinência de estudos de previsão de carga. Segue-se um enquadramento relativo ao tema da tese, onde se refere a produção elétrica nacional na atualidade, face a outras formas de produção de energia. Além disso, são abordados os objetivos para o desenvolvimento da dissertação e uma breve descrição da sua estrutura.

1.1 Considerações Gerais

Em vários países do mundo, o setor de energia elétrica tem vindo a sofrer um processo de transformação, visto que em muitos deles o modelo em que se estruturava o setor elétrico era em monopólio vertical e controlado pelos respetivos governos, ou noutros casos a estrutura do setor elétrico dependia de uma única empresa monopolista. Com a privatização e a desverticalização do setor elétrico, deu-se uma mudança de cenário onde a eletricidade pôde ser livremente negociada e comercializada. Desta forma, o sistema de energia elétrica evoluiu na direção de estabelecer uma estrutura de mercado, constituindo os diferentes setores envolvidos no mercado de energia elétrica, independentes e autónomos entre si: geração, transmissão, distribuição e comercialização.

Com a reestruturação do sistema elétrico e a liberalização do mercado de energia, as diferentes empresas envolvidas no setor, promovem a competitividade económica do mercado elétrico, fornecendo ao consumidor final um serviço de qualidade com custos competitivos. Devido a estas alterações no sistema elétrico, percebe-se que a realização de estudos sobre previsão de carga, assume bastante interesse numa perspetiva económica, pois os erros de previsão podem conduzir a custos avultados para as empresas do mercado elétrico. Estes erros associados à previsão de carga, podem ser por excesso (*overforecast*) ou por defeito (*underforecast*), face à carga verificada/real. Erros por excesso significam custos

adicionais desnecessários, uma vez que podem levar à ligação de geradores adicionais que seriam dispensáveis. Erros por defeito são habitualmente ainda mais graves, dado que podem levar a insuficiência de geração para satisfazer toda a carga.

Relativamente à previsão de carga elétrica num determinado intervalo de tempo, normalmente distingue-se em três categorias:

- de curto prazo, de alguns minutos até uma semana à frente;
- de médio prazo, de uma semana a um ano;
- de longo prazo, superiores a um ano.

Com o objetivo de otimizar as tarefas de um sistema de energia de grande escala, diversas técnicas podem ser aplicadas com uma considerável expectativa de economia nos custos e de eficiência do sistema, daí que vários estudos têm sido desenvolvidos nesta área.

A qualidade da projeção futura do consumo de eletricidade a ser realizada pelas distribuidoras de energia, afeta significativamente dois pontos distintos. O primeiro refere-se ao planeamento da expansão do sistema elétrico, uma vez que estas previsões são utilizadas para dar suporte às decisões tomadas neste sentido. O segundo ponto está no contexto operacional e financeiro das próprias distribuidoras. Qualquer erro que ultrapasse os limites estabelecidos, acarreta em perdas e afeta diretamente os resultados financeiros.

Tendo em vista o contexto criado pelas mudanças estruturais do setor, é necessário por parte das distribuidoras e comercializadores de energia, investir em técnicas e metodologias para gerir as incertezas inerentes ao mercado. As tomadas de decisão relativas à contratação futura e a própria gestão dos contratos (em escala anual ou mensal) devem estar comprovados por previsões satisfatórias acerca do consumo de energia.

O desafio de prever o consumo de energia elétrica no contexto de uma economia estável, com séries temporais razoavelmente estacionárias, é por si só um desafio. Esse desafio torna-se tanto maior no contexto de uma economia cujas características mudam significativamente em curtos intervalos de tempo.

Conforme referido, a dinâmica do consumo de energia possui uma complexidade inerente ao seu processo. A influência de variáveis económicas e climáticas, juntamente com a dificuldade de se conhecer as entradas do sistema e a relação destas com o consumo de energia, torna difícil o uso de modelos analíticos para tratar o problema. Por outro lado, o uso de técnicas baseadas em identificação de sistemas e inteligência computacional (como redes neurais e SVM) têm sido muito utilizadas em diversos estudos realizados com sucesso.

1.2 Enquadramento

A energia elétrica desempenha um papel relevante, tanto na economia como no bem-estar e comodidade da sociedade atual. É expectável que um sistema de energia produza eletricidade da forma mais económica possível e com um alto nível de qualidade e fiabilidade. Tendo em conta a imensa dependência energética global, é necessário garantir a satisfação da procura de energia elétrica e para isso, é crucial que haja uma boa gestão do sistema de energia, desde a produção, ao transporte e consumo, de modo a que os vários limites técnicos do sistema sejam cumpridos e assim, se garantir um funcionamento adequado.

A título de exemplo, observa-se na figura 1.1 o diagrama de carga do dia 24 de novembro de 2012 com os traçados referentes à carga prevista e à verificada, onde se visualiza a variação da potência necessária para satisfazer o consumo nacional ao longo deste dia.

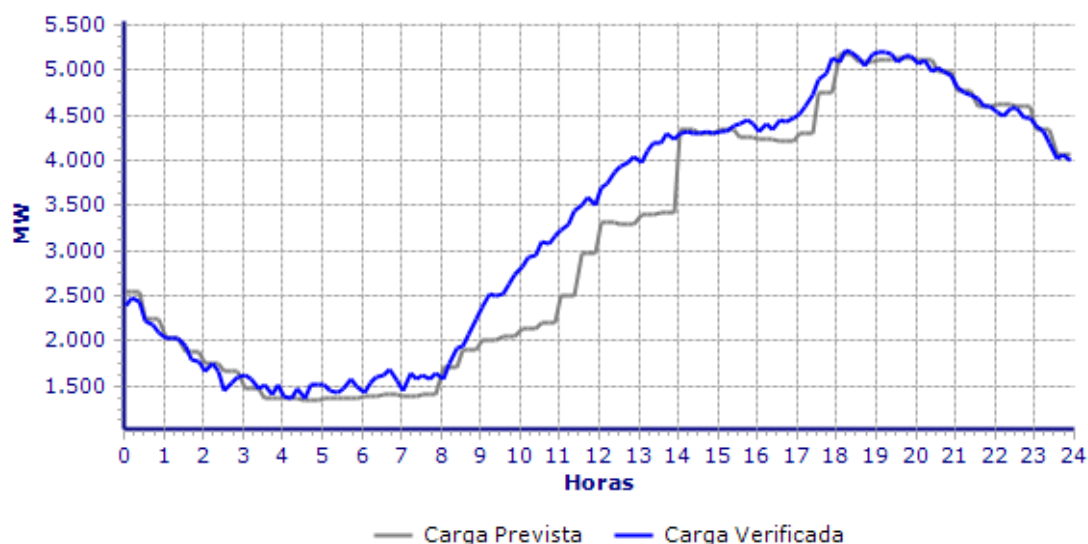


Figura 1.1 – Diagrama de carga do dia 24 de novembro de 2012 [1].

Neste exemplo da figura 1.1, percebe-se que existe alguma diferença da carga prevista para a carga verificada, nomeadamente entre as 9 e as 14 horas. Verifica-se com este exemplo, que esta previsão apesar de ter sido supostamente feita por profissionais apresenta ocasionalmente erros consideráveis. Tendo em conta esta observação, é de salientar a importância do estudo das diferentes técnicas de previsão para determinar qual a mais precisa e a que mais se possa aproximar à carga a verificar.

Na figura 1.2, observa-se o diagrama de cargas do dia 9 de outubro de 2013. Com este exemplo pretende-se destacar os principais tipos de produção de energia existentes no diagrama de carga nacional, onde na sua base nota-se a produção das centrais térmicas (carvão), seguida da produção em regime especial (PRE) que engloba

várias fontes de energia renovável. Torna-se evidente que a dependência energética face ao exterior assume um papel de relevo no diagrama de carga, o que resulta em custos preponderantes para os interesses económicos nacionais.

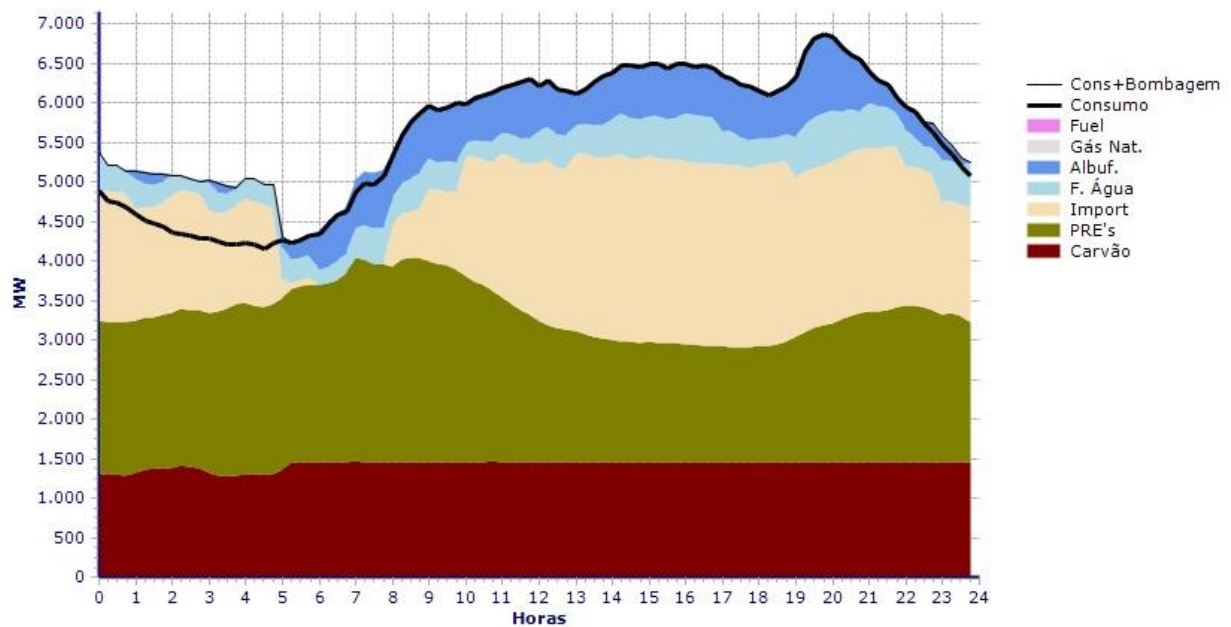


Figura 1.2 – Diagrama de carga do dia 9 de outubro de 2013 [1].

Visto que a PRE assume uma importância considerável no diagrama de carga nacional, salienta-se a aposta e o aumento que tem havido na produção com fontes de energia renováveis, visível na figura 1.3. Este facto reforça novamente a necessidade de previsões de carga de elevada qualidade, de modo a contribuir de forma contínua para o aumento sustentado da integração de produção renovável.

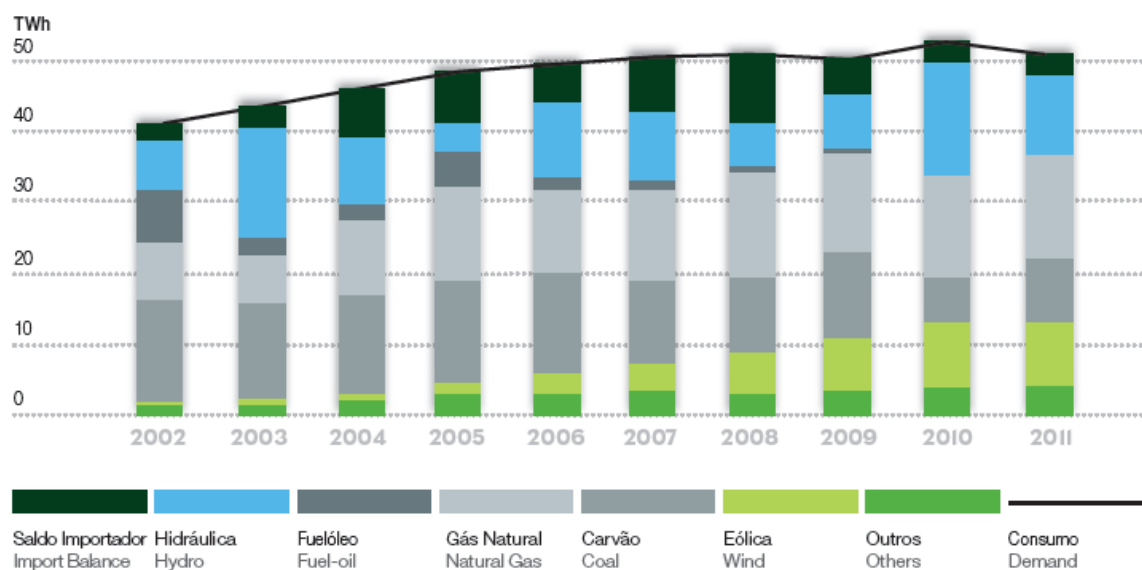


Figura 1.3 – Fontes de produção de energia em Portugal [2].

1.3 Objetivos

Esta dissertação tem como objetivo principal comparar o desempenho de diferentes técnicas de previsão de carga, especialmente a curto prazo. Inicialmente pretende-se realizar a análise clássica de séries temporais, obtendo assim uma referência a nível de performance. Com os resultados obtidos deste primeiro estudo, espera-se que estes sejam comparados e ultrapassados pelos das outras duas técnicas (redes neuronais e SVM). Esta análise de resultados será feita por avaliação de diferentes tipos de índices (de erros), assim como tempos de treino, flexibilidade e capacidade de filtragem de dados anómalos.

1.4 Estrutura da Dissertação

A dissertação encontra-se estruturada em sete capítulos principais, cujo conteúdo será, resumidamente, descrito de seguida.

No Capítulo 1 “Introdução”, aborda-se de uma forma geral a estrutura do sistema elétrico nacional e a relevância de estudos de previsão de carga. Além disso, é elaborado um enquadramento do tema da tese ao diagrama de carga nacional, são referidos os objetivos da dissertação e é descrita a sua estrutura.

No Capítulo 2 “Métodos de Previsão”, pretende-se expor conteúdos relativos às técnicas utilizadas para a previsão de carga, onde se explicam os conceitos associados a cada uma. Além disso, definem-se as equações de cálculo que permitem a avaliação do desempenho dos modelos em estudo.

No Capítulo 3 “Previsão baseada em Box-Jenkins”, apresentam-se os procedimentos referentes à metodologia em estudo, assim como os respetivos resultados. No final do capítulo, comparam-se os modelos estudados através dos critérios que avaliam o seu desempenho, para assim se apurar o melhor e obter uma referência para os próximos capítulos.

No Capítulo 4 “Previsão com Redes Neuronais”, expõem-se os resultados obtidos com o estudo efetuado, referindo os passos mais relevantes da utilização desta técnica.

No Capítulo 5 “Previsão com SVM”, apresentam-se os resultados obtidos com a aplicação técnica ao problema de previsão.

No Capítulo 6 “Comparação das Técnicas”, comparam-se os resultados obtidos com o estudo das técnicas utilizadas para a previsão de carga e referem-se as conclusões em relação ao desempenho de cada uma.

No Capítulo 7 “Conclusão”, são sintetizadas as conclusões mais relevantes obtidas com o desenvolvimento do trabalho. Sugestões para futuros desenvolvimentos são também apresentadas neste capítulo.

1.5 Software Utilizado

Como ferramentas informáticas de auxílio ao desenvolvimento da dissertação, foram utilizados os programas *Microsoft Excel*, *Mathworks Matlab* e *SPSS Statistics*. Numa primeira fase do trabalho, ao conjunto de dados fornecidos em Excel, foi feita uma filtragem e separação por anos, além da criação de gráficos de carga anuais, por forma a se perceber a variação de carga nos diferentes anos. O SPSS foi utilizado para o modelo Box-Jenkins, tendo sido o Matlab utilizado para realização de estudos relativos a redes neuronais e a SVM.

Capítulo 2

Métodos de Previsão

Neste capítulo, expõem-se conteúdos relativos às técnicas utilizadas para a previsão de carga, onde se explicam os conceitos associados a cada uma. Posteriormente, referem-se quais as equações/critérios que são utilizados neste trabalho, que permitem avaliar o desempenho dos modelos em estudo.

2.1 Modelo Clássico (Box-Jenkins)

Para a previsão de séries temporais, o modelo clássico que se tornou mais popular é o designado por Box-Jenkins. Este modelo pretende prever valores futuros de uma série, baseando-se nos seus valores passados. Outra designação para a metodologia de previsão Box-Jenkins, é conhecida por modelos ARIMA (AR: auto-regressivos, I: integrados, MA: média móvel) [3].

A metodologia Box-Jenkins começa pela tentativa de identificação do modelo, isto é, a determinação dos três componentes que configuram qualquer estrutura ARIMA (p, d, q):

- p - parâmetros auto-regressivos;
- d - processos de diferenciação (integração);
- q - parâmetros de médias móveis.

Por exemplo, a estrutura ARIMA (1, 2, 0) significa que esta possui um parâmetro auto-regressivo, duas diferenciações a partir da série original e nenhum parâmetro de médias móveis. O procedimento padrão para utilização da metodologia de previsão ARIMA consiste nos seguintes passos [3, 4, 5]:

1. Realiza-se a diferenciação da série original tantas vezes quanto necessário para torná-la estacionária (média e variância constantes);

O primeiro passo desta metodologia consiste em verificar a estacionariedade da série. Caso se verifique que não é estacionária, deve ser sucessivamente diferenciada, até que esta condição se torne válida. O número de diferenças necessárias, é denominado ordem de integração (I).

Diferenciação de primeira ordem para uma primeira diferença:

$$X'_t = (1-B)X_t = X_t - X_{t-1} \quad (2.1)$$

Onde “B” é o operador “Back”. Exemplo: $B \cdot X(t) = X_{t-1}$.

Diferenciação de primeira ordem para uma segunda diferença:

$$X'_t = (1-B^2)X_t = X_t - X_{t-2} \quad (2.2)$$

Diferenciação de segunda ordem:

$$X'_t = (1-B)^2 X_t = (X_t - X_{t-1}) - (X_{t-1} - X_{t-2}) \quad (2.3)$$

Em alternativa à diferenciação, a série pode ser sujeita a algum tipo de transformação (por exemplo logarítmica, exponencial, entre outras), ou então calculando a média deslizante e o desvio padrão deslizante em janelas de 30 dias (como se observará no capítulo seguinte [3]).

2. Identificam-se os valores p e q através da análise das funções de auto-correlação (ACF) e de auto-correlação parcial (PACF) estimadas. A análise destas funções permite a tomada de decisão acerca do número p de parâmetros auto-regressivos e do número q de parâmetros de médias móveis, que devem ser escolhidos de maneira a se obter uma estrutura que se ajuste aos dados da série. De todas as hipóteses possíveis, deve ser selecionada aquela que tenha o menor número de parâmetros;
3. Os dois passos anteriores permitem identificar o modelo que, em termos gerais, assumirá uma forma do tipo:

$$\phi(B)(1-B)^d X_t = \theta(B)a_t \quad (2.4)$$

em que:

$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ é o operador auto-regressivo (AR) de ordem p.

$(1 - B)^d$ é o operador de integração (I) de ordem d.

$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$ é o operador de médias móveis (MA) de ordem q.

a_t corresponde ao erro gerado pelo modelo.

Uma vez estabelecido o modelo (equação de referência), passa-se à determinação dos coeficientes desta equação, ou seja, procede-se à estimação dos parâmetros (ϕ_i , $i = 1, \dots, p$ e θ_i , $i = 1, \dots, q$).

4. O passo seguinte inclui a realização de testes com o modelo determinado no passo anterior e determinação do erro por comparação da previsão com o valor real. Em seguida estimam-se os valores ACF e PACF para a série do erro. Estes indicadores ditarão a necessidade ou não de incluir mais termos no modelo previamente estabelecido.

No caso de a série temporal em estudo poder apresentar algum componente sazonal, o modelo é o designado por SARIMA (*Seasonal Auto Regressive Integrated Moving Average*). Este baseia-se na equação 2.4 referida para os modelos ARIMA, acrescentando os respetivos componentes sazonais. O seu modelo geral é dado pela equação 2.5.

$$(1-B)^d(1-\phi_1B-\phi_2B^2-\dots-\phi_pB^p)(1-\Phi_1B^P)(1-\Phi_2B^{2P}-\dots-\Phi_qB^{qP})X_t = (1-\theta_1B-\theta_2B^2-\dots-\theta_qB^q)(1-\Theta_1B^Q)(1-\Theta_2B^{2Q}-\dots-\Theta_rB^{rQ})a_t \quad (2.5)$$

Onde:

- $(1-\Phi_1B^P)$ é o componente sazonal AR de ordem P;
- $(1-\Theta_1B^Q)$ é o componente sazonal MA de ordem Q.

Resumidamente, para a construção de um modelo ARIMA, os procedimentos a seguir são:

1. observar gráficos, identificar estacionaridade, identificar sazonalidade;
2. caso necessário (série não estacionária), diferenciar a série ou aplicar transformações logarítmicas, exponenciais ou potências de forma a garantir a sua estacionaridade;
3. calcular ACF e PACF para identificar o modelo;
4. determinar os parâmetros do modelo;
5. realizar testes: fazer a previsão, validar o modelo, avaliar o desempenho do modelo;
6. calcular ACF e PACF do modelo. Caso o erro revele ainda termos ACF ou PACF importantes, integra-los no modelo e voltar ao passo 4.

A figura 2.1 apresenta um fluxograma que sintetiza a técnica usada para construção de um modelo ARIMA.

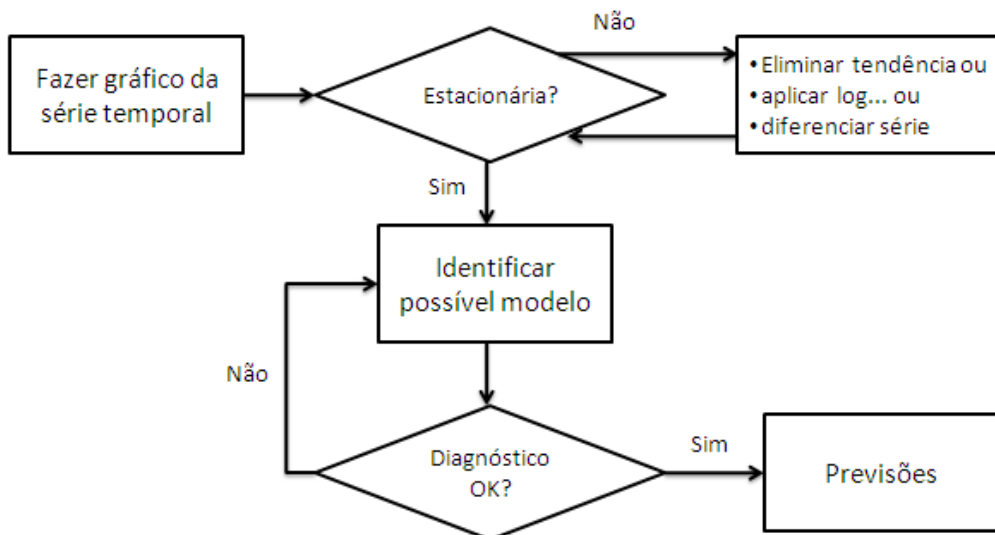


Figura 2.1 – Fluxograma para a construção de um modelo ARIMA.

2.2 Redes Neurais

Uma rede neuronal consiste numa estrutura, na qual o processamento se encontra distribuído por um grande número de unidades relativamente simples mas densamente interligadas. Por analogia com os sistemas biológicos, estas unidades processadoras são usualmente designadas por neurónios.

Um neurónio biológico, consiste numa única célula capaz de realizar uma forma simples de processamento. Cada neurónio é estimulado por uma ou mais ligações vindas de outros neurónios, chamadas sinapses, dependendo o sinal produzido tanto da força das ligações como da sua natureza. Este sinal é propagado ao longo do axónio, indo por sua vez, estimular outros neurónios. O funcionamento dos neurónios artificiais¹ baseia-se, na generalidade dos casos, neste modelo simplificado dos neurónios biológicos.

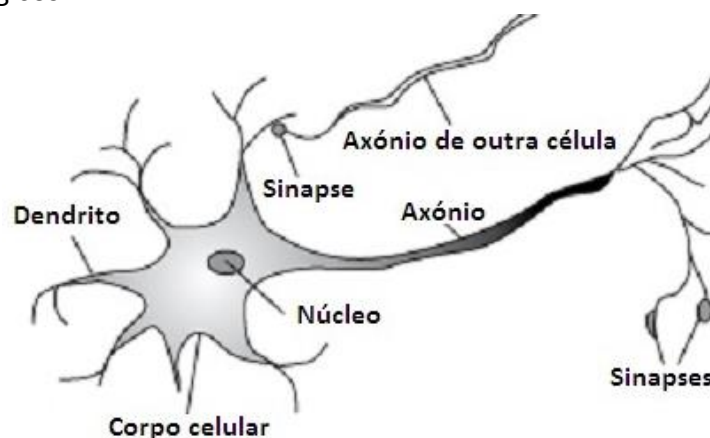


Figura 2.2 – Constituição de um neurónio biológico [6].

¹ Os neurónios artificiais são frequentemente designados apenas por “unidades”.

As redes neuronais artificiais (RN) são técnicas computacionais que apresentam um modelo matemático inspirado na estrutura dos neurónios reais referidos anteriormente, que adquirem conhecimento através da experiência.

Estas redes são constituídas por várias unidades de processamento conectadas por canais de comunicação, cada um dos quais com um determinado peso.

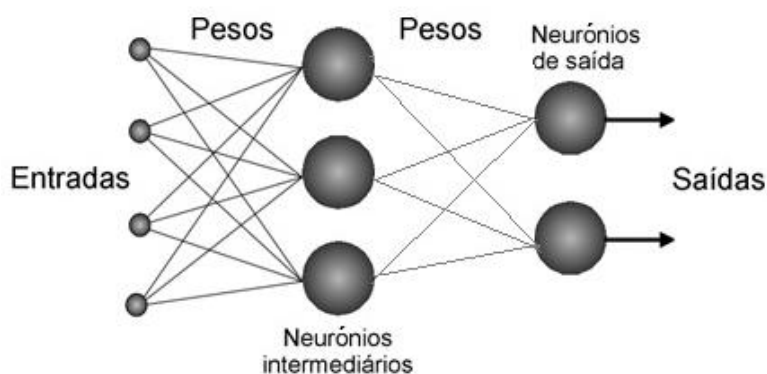


Figura 2.3 – Esquema exemplificativo de uma rede neuronal.

A unidade básica de uma RN é designada por neurónio artificial. Este considera as várias entradas possíveis, processando a informação numérica fornecida e os respetivos pesos associados, resultando assim numa resposta.

De forma a se perceber melhor o modelo matemático de um neurónio artificial, apresenta-se a figura 2.4.

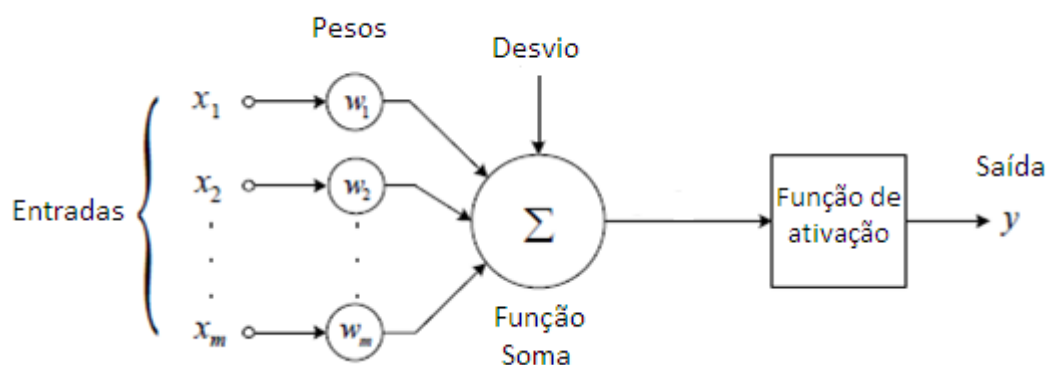


Figura 2.4 – Modelo genérico de um neurónio artificial.

Através da observação da figura 2.4, é possível identificar os elementos básicos de um neurónio artificial, nomeadamente as entradas “ x_i ”, os pesos das ligações “ w_i ”, uma função soma, uma função de ativação e a saída “ y ”.

É através da função de ativação que é possível a introdução de componentes não lineares na RN. Estas podem ser de diferentes tipos, pelo que se mencionam duas das mais utilizadas, na figura 2.5.

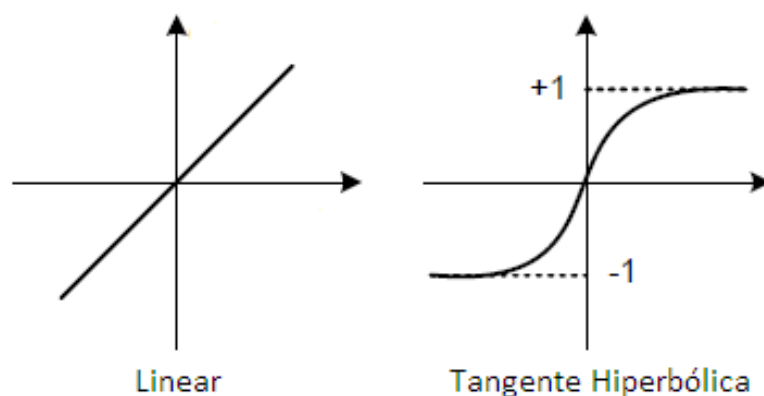


Figura 2.5 – Tipos de funções de ativação típicas em RN.

Em relação à função de ativação linear, os valores de saída desta função correspondem a toda a gama do conjunto real, existindo correspondência direta entre as entradas e as saídas. No caso da função tangente hiperbólica, os valores de saída podem variar entre -1 e 1.

As RN têm sido usadas com sucesso em diversas aplicações tais como aproximação de funções, classificação de padrões, análise de regressões e previsão.

Um aspeto importante para o trabalho com redes neuronais são os algoritmos de treino que podem ser classificados em dois grandes grupos: a aprendizagem supervisionada ou a não supervisionada. A aprendizagem supervisionada (a usada neste trabalho) é caracterizada pelo facto de a RN ser treinada com exemplos entrada-saída. A título de exemplo, enumeram-se alguns algoritmos de treino de aprendizagem supervisionada, utilizados em RN: regra de aprendizagem do perceptron, algoritmo de retro-propagação do erro (*Backpropagation* na literatura anglo-saxónica), método de Hebb, algoritmo Levenberg-Marquardt, algoritmo do gradiente conjugado, entre outros [7, 8, 9].

Quanto à direção do fluxo de informação, as RN podem ser classificadas em dois tipos: unidireccionais (*feedforward*) ou recorrentes (*feedback*). Nas unidireccionais, o fluxo de informação é sempre na direção da entrada para a saída. Nas recorrentes, existem ligações de unidades mais próximas da saída para unidades mais próximas da entrada (normalmente da própria saída para as unidades de entrada).

Uma das características principais das RN é a de possuírem capacidade de adaptação à evolução do conjunto de dados, o que permite efetuar treinos a qualquer momento, permitindo uma atualização constante da RN. Uma outra forma de agir é com uma aprendizagem da RN do tipo adaptativo, que tem a possibilidade da rede aprender sobre o sistema em tempo real [7, 8].

Por outro lado, as RN necessitam de dispor de um grande conjunto de dados, para que seja possível proceder ao processo de aprendizagem da rede e precisa ainda de um elevado poder de processamento computacional, no caso de a rede ser composta por um elevado número de entradas e de neurónios. Outro aspeto relevante

relaciona-se com a possibilidade do processo de treino convergir para um mínimo local da função de minimização, o que pode conduzir a um erro significativo. Neste caso, convém reiniciar o treino, uma vez que os pesos são inicializados de forma aleatória, a probabilidade da RN convergir para o mesmo mínimo local é ínfima.

É fundamental ter em conta que utilizar um número de neurónios ocultos maior do que o necessário, pode resultar na perda da capacidade de generalização da RN, o que se designa por *overfitting*.

A utilização de RN em sistemas de previsão, consiste no facto de não ser necessário saber a natureza da série de dados, uma vez que o processo de aprendizagem obtém as características essenciais do conjunto de dados.

Para que seja possível obter um bom modelo de previsão da RN em estudo, devem ser consideradas as seguintes etapas:

- Definir o problema identificando as variáveis a prever assim como as variáveis independentes relevantes;
- Separar, filtrar e processar os dados como entradas e saídas, para treinar e testar a rede;
- Treinar a rede e avaliar o desempenho;
- Testar de novo a rede caso o desempenho do treino anterior não ser satisfatório;
- Obtenção dos valores da previsão.

2.3 Máquinas de Suporte Vetorial (SVM)

As máquinas de suporte vetorial (SVM: *support vector machines*) são modelos de aprendizagem supervisionada com a associação de algoritmos de aprendizagem, inicialmente criados para solução de problemas de classificação [10]. Ao contrário de métodos estatísticos clássicos aplicados à classificação, que precisam de uma quantidade elevada de dados juntamente com o conhecimento prévio do problema abordado, este tipo de aprendizagem supervisionada foi desenvolvido para a resolução de problemas de classificação cuja quantidade de dados disponíveis é reduzida e o conhecimento prévio sobre o problema ser pouco ou mesmo inexistente, características estas encontradas em muitas aplicações reais.

Os modelos SVM aplicam-se por exemplo, na análise de dados e reconhecimento de padrões, utilizados para classificação e análise de regressão. O SVM convencional, considera como entrada um conjunto de dados e prevê para cada entrada considerada, qual de duas possíveis classes a entrada faz parte, o que faz do SVM um classificador linear binário não probabilístico. Dado um conjunto de exemplos de treino, cada um marcado como pertencente a uma de duas categorias, um algoritmo de treino do SVM constrói um modelo que atribui novos exemplos a uma

categoria ou outra. Um modelo SVM é uma representação de exemplos como pontos no espaço, mapeados de maneira a que os exemplos de cada categoria sejam divididos por um espaço definido que seja tão amplo quanto possível. Os novos exemplos são então mapeados no mesmo espaço e previstos como pertencentes a uma das categorias, onde são depois colocados.

A teoria das SVM foi originalmente elaborada para a solução de problemas de classificação, através da aplicação do conceito de hiperplano ótimo de separação, baseado na maximização da margem de separação (ρ). A figura 2.4 ilustra a margem de separação “ ρ ” para o caso de padrões linearmente separáveis em duas classes. Nesta figura, a reta vermelha representa o hiperplano ótimo de separação, com os chamados vetores de suporte sendo aqueles situados exatamente em cima das retas a tracejado. Estes vetores recebem esta denominação em virtude da sua proximidade da superfície de decisão, contribuindo de maneira importante para a definição de tal superfície [10].

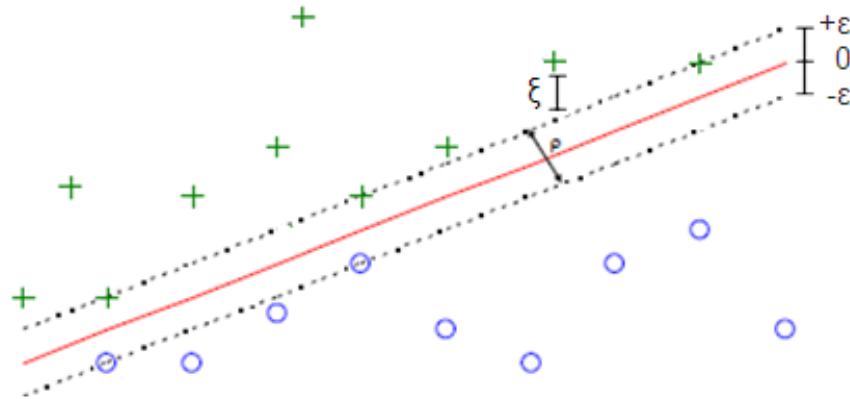


Figura 2.6 – Margem de separação ρ para o caso de duas classes linearmente separáveis.

Aplicando as SVM ao desenvolvimento deste trabalho, pretende-se aplicar o modelo de regressão de suporte vetorial (SVR) que permite a previsão de séries temporais. O SVR resolve um problema de otimização (neste caso de minimização), segundo as seguintes equações:

Função objetivo:

$$\min (w, b, \xi, \xi^*): \frac{1}{2} ||w||^2 + C * \sum_{i=1}^l (\xi_i + \xi_i^*) \quad (2.6)$$

Sujeito às restrições:

$$y_i - (w^T \cdot \phi(x_i) + b) \leq \epsilon + \xi_i \quad (2.7)$$

$$(w^T \cdot \phi(x_i) + b) - y_i \leq \epsilon + \xi_i^* \quad (2.8)$$

com $\xi_i, \xi_i^* \geq 0, i = 1, \dots, l$

A variável “xi” representa os dados em estudo da série de consumo, mapeada num espaço dimensional pela função ϕ , onde “w” é o vetor normal ao hiperplano e “b” é o desvio, ξ_i é o erro superior de treino relativamente ao hiperplano ótimo de separação, enquanto ξ_i^* é o inferior, sujeitos a:

$$|y - (w^T \cdot \phi(x) + b)| \leq \varepsilon$$

Os parâmetros que controlam a qualidade da regressão são: “C” que é o fator de penalização do erro, a largura entre o hiperplano ótimo de separação para os vetores de suporte “ ε ” e a função de mapeamento ϕ . A variável ε , define o valor da tolerância ao erro e o seu grau de precisão será o especificado, tendo em conta a regressão a efetuar.

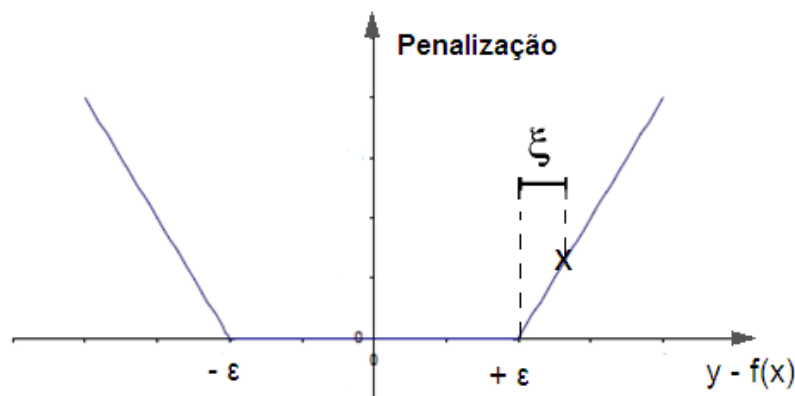


Figura 2.7 – Gráfico da função de penalização para um determinado ε .

As restrições da equação 2.6 e 2.7 implicam que se tente abranger o máximo possível de dados “xi”, no entanto aos que ficam no exterior do espaço definido pelos vetores de suporte, considera-se a função de penalização dada por:

$$|y - (w^T \cdot \phi(x) + b)| \leq \varepsilon \quad (2.9)$$

Se algum vetor não estiver dentro do espaço definido pelos vetores de suporte, então existe um erro ξ_i ou ξ_i^* que se pretende minimizar pela função objetivo. Por observação da figura 2.7 percebe-se a relação entre uma possível penalização a um vetor que está situado no exterior, do espaço definido pelos vetores de suporte.

Para a realização deste trabalho, foi utilizado o *software* Matlab R2014a através da utilização da *toolbox* de SVM desenvolvida por *Steve Gunn*².

² Disponível em: <http://www.isis.ecs.soton.ac.uk/resources/svminfo>

Comparando os modelos SVM com as RN, a principal vantagem que tem sido apontada aos primeiros, é o seu conceito de minimização de uma tolerância máxima para o erro de generalização, ao contrário do processo de treino das RN, que simplesmente permite minimizar o erro associado ao conjunto de dados segundo os quais a rede é treinada [10]. Desta forma, as SVM garantem uma solução única e ótima e por esse motivo, é que esta técnica é apontada como uma evolução das RN, tanto pelos acertos da previsão, como pelo processo acessível de calibração de parâmetros durante a fase de treino.

2.4 Avaliação do desempenho dos modelos

A eficácia da previsão dos modelos estudados, verifica-se através da avaliação ao seu desempenho. Para tal, definem-se as equações do erro médio quadrático (RMSE), calculado pela média dos quadrados dos erros de previsão e do erro médio percentual absoluto (MAPE), que representa a média percentual da divisão entre erro absoluto de previsão e o valor real [11].

Erro médio quadrático (RMSE):

$$RMSE = \sqrt{\frac{1}{n} \cdot \sum_{t=1}^n (Y_t - \hat{Y}_t)^2} \quad (2.10)$$

Erro médio percentual absoluto (MAPE):

$$MAPE = \frac{1}{n} \cdot \sum_{t=1}^n \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right| \quad (2.11)$$

Em que Y_t é o valor real da série, \hat{Y}_t é o valor previsto e “n” é o número de observações da previsão.

Estas fórmulas são úteis para a comparação entre modelos, quando o critério de decisão for escolher qual o que fornece a melhor previsão. Os modelos com menores valores de RMSE e MAPE ou seja, com os menores erros, são os que fornecem as previsões mais eficazes.

2.5 Conclusão

Neste capítulo apresentaram-se os conceitos base relativos às diferentes técnicas utilizadas nos estudos de previsão de carga, nomeadamente: Box-Jenkins, RN e SVM. Além disso, foram também definidos os critérios de avaliação do desempenho dos modelos em estudo, tendo-se optado pelos indicadores RMSE e MAPE.

Capítulo 3

Previsão baseada em Box-Jenkins

No capítulo 2 foram apresentadas as três técnicas que são utilizadas para os estudos de previsão, neste trabalho. Desta forma, neste novo capítulo pretende-se iniciar o estudo de séries temporais utilizando a análise clássica Box-Jenkins, considerando como horizonte de previsão a hora seguinte.

3.1 Verificação da condição de estacionariedade

Considerando os procedimentos respetivos à sua metodologia, pretende-se traçar o gráfico da série de consumos em estudo, de forma a observar a sua condição de estacionariedade.

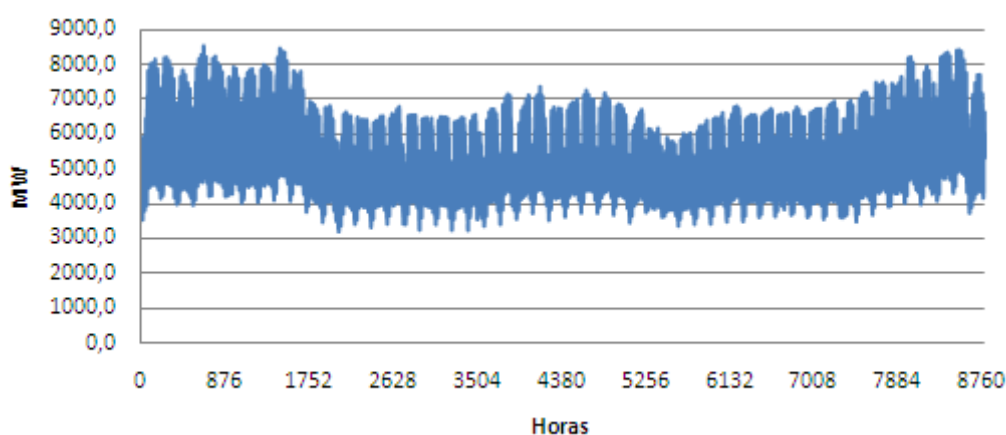


Figura 3.1 – Série relativa ao consumo do ano de 2005.

Observando o gráfico da figura 3.1 da série de consumo em estudo, nota-se que existe um padrão de variação da média e do desvio padrão ao longo do ano, pelo que se conclui que esta série não é estacionária. Segundo a metodologia Box-Jenkins para o estudo de séries temporais, é condição obrigatória que a série em estudo seja estacionária. Assim sendo, observa-se a média e o desvio padrão numa janela mensal de 30 dias, aplicadas à série de consumo.

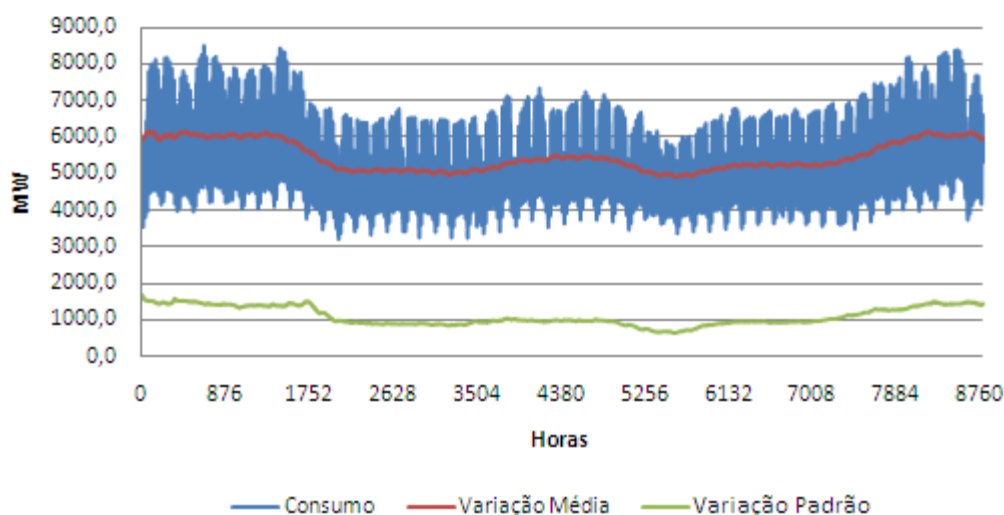


Figura 3.2 – Série de consumo, com a média e o desvio padrão em janelas de 30 dias.

Por observação da figura 3.2, nota-se que tanto a média como o desvio padrão determinados, variam ao longo do ano seguindo um certo padrão. Além disso, para obter a série estacionária será necessário calcular o desvio padrão de toda a série. Assim sendo, considera-se a equação 3.1 para a determinação da série estacionária [3].

$$C = (\text{Consumo} - \text{Cons. Médio}_{30}) * \text{Var. Padrão} / \text{Var}_{30} \quad (3.1)$$

Onde:

- Consumo refere-se à série em estudo;
- Cons. Médio₃₀ refere-se à média deslizando aplicada em janelas de 30 dias da série de consumo;
- Var. Padrão é o desvio padrão de toda a série de consumo;
- Var₃₀ refere-se ao desvio padrão deslizando aplicado em janelas de 30 dias da série de consumo.

Tendo realizado os cálculos no Excel, observa-se na figura 3.3 a série estacionária.

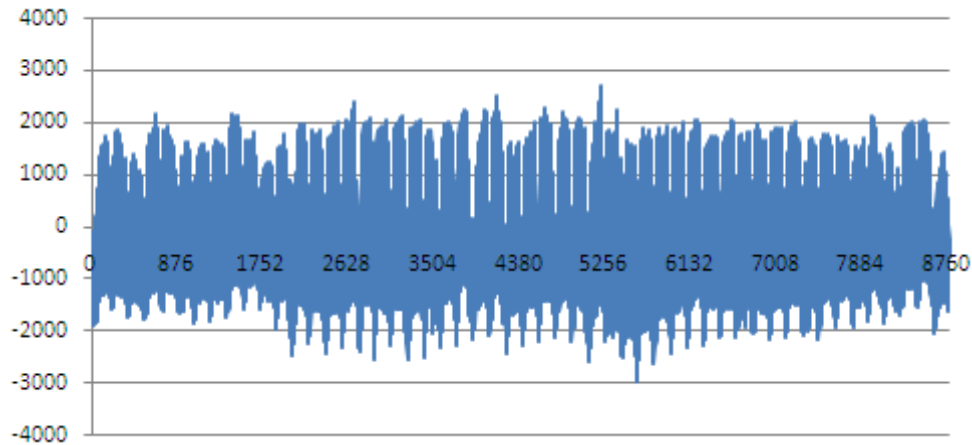


Figura 3.3 – Série de consumo estacionária.

De referir que para recuperar a série de consumo original a partir da estacionária, será necessário utilizar valores aproximados para o consumo médio e o desvio padrão. Através da equação 3.2 consegue-se voltar à série original, a partir da estacionária.

$$\text{Consumo} = (C * \text{Var.}/\text{Var. Padrão}) + \text{Cons. Médio} \quad (3.2)$$

3.2 Identificação do modelo

Com a série estacionária definida, pretende-se agora calcular o ACF e PACF de forma a identificar o modelo para o estudo. Para este efeito, utilizou-se o *software* SPSS definindo um *lag* de 170 (visto que uma semana tem 168 horas).

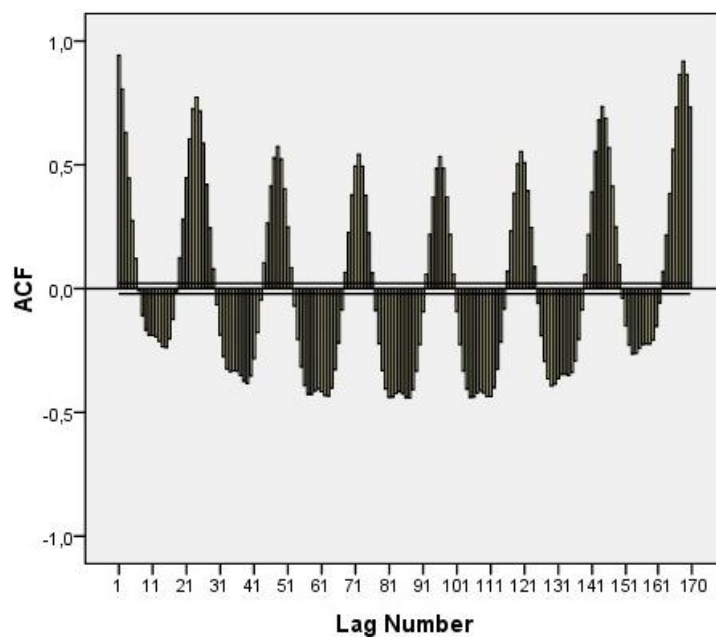


Figura 3.4 – Gráfico ACF da série estacionária.

Na figura 3.4 observa-se o ACF onde se nota um comportamento sinusoidal atenuado, típico do processo auto-regressivo. Nota-se também que os termos correspondentes a uma semana antes são bastante fortes, o que indica uma forte correlação da carga no dia “D” com a carga no dia “D-1” semana.

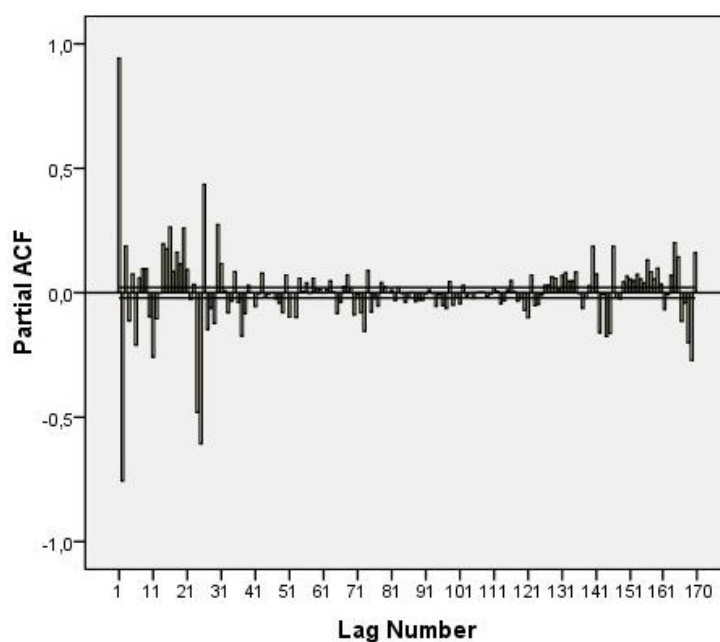


Figura 3.5 – Gráfico PACF da série estacionária.

Em relação à figura 3.5, observa-se o PACF onde se destacam as 2 primeiras barras, o que reforça a ideia de um modelo auto-regressivo, neste caso um AR(2).

3.3 Estimação de parâmetros

Após a identificação do modelo, pretende-se agora estimar os seus parâmetros. Sabe-se que os componentes do modelo identificado são: $p = 2$, $q = 0$. Ou seja, substitui-se o operador AR(2) por $\phi(B) = 1 - \phi_1 B - \phi_2 B^2$ e quanto ao operador de MA(0) substitui-se por $\theta(B) = 1$, resultando a equação 3.3. Desenvolvendo-a, obtêm-se as equações 3.4 e 3.5.

$$(1 - \phi_1 B - \phi_2 B^2)(1-B)X_t = a_t \quad (3.3)$$

$$(1 - \phi_1 B - \phi_2 B^2 - B + \phi_1 B^2 + \phi_2 B^3)X_t = a_t \quad (3.4)$$

$$X_t - \phi_1 B^* X_t - \phi_2 B^2 X_t - B^* X_t + \phi_1 B^2 X_t + \phi_2 B^3 X_t = a_t \quad (3.5)$$

Escrevendo a equação anterior em ordem a X_t e substituindo o operador B , tendo em conta o exemplo: $B^* X(t) = X_{t-1}$, resulta a equação 3.6:

$$X_t = (1 + \phi_1)X_{t-1} + (\phi_2 - \phi_1)X_{t-2} - \phi_2 X_{t-3} + a_t \quad (3.6)$$

A equação 3.6 apresenta o modelo desenvolvido. De forma a estimar os parâmetros do modelo proposto, será utilizado o método dos mínimos quadrados com o “solver” do Excel. Na estimação dos coeficientes, considera-se que $a_t = 0$.

Os valores obtidos foram: $\phi_1 = 0,938$ e $\phi_2 = -0,327$ para os coeficientes e de $5,8674E+08$ para o somatório dos erros ao quadrado ($\sum a_t^2$). Deste modo, substituindo os valores na equação anterior, apresenta-se a equação 3.7 relativa ao modelo estimado X_t^{\wedge} .

$$X_t^{\wedge} = 1,938 * X_{t-1} - 1,265 * X_{t-2} + 0,327 * X_{t-3} + a_t \quad (3.7)$$

3.4 Comparação do consumo real com o estimado

Com o modelo estimado, observam-se as figuras 3.6 e 3.7 onde se apresenta a dispersão entre a série de consumo e a estimada (com um coeficiente de correlação R^2 de 0,9607) e uma amostra de 340 horas com a comparação das séries, respetivamente.

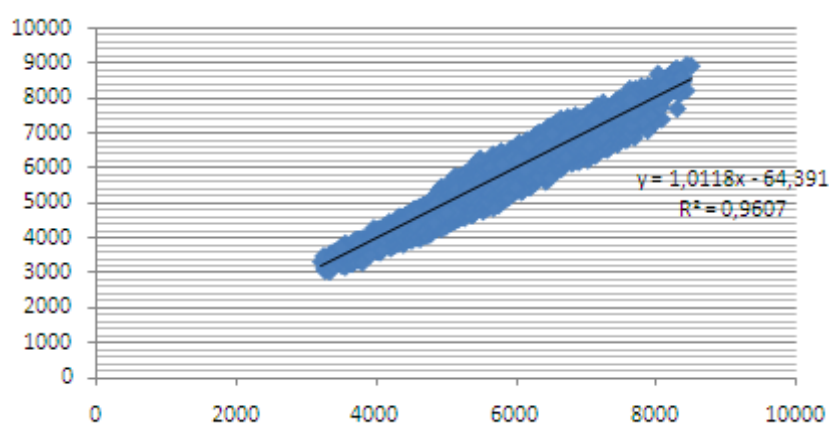


Figura 3.6 – Correlograma do Consumo Real Vs Estimado.

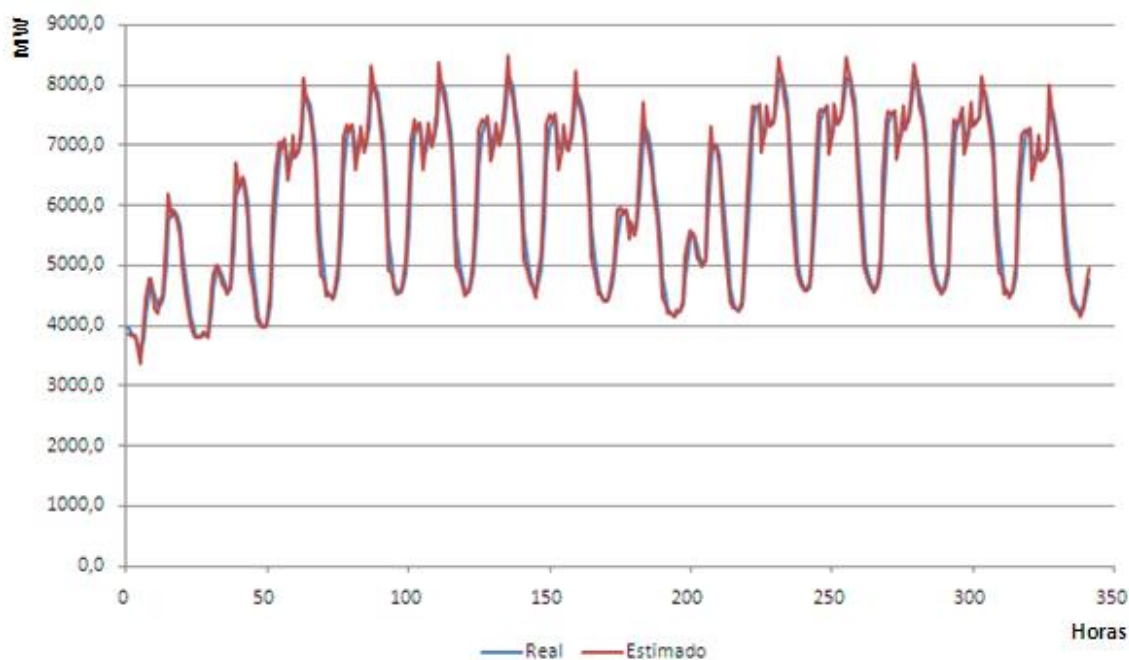


Figura 3.7 – Comparação do Consumo Real com o Estimado.

Na figura 3.7 compara-se uma parte da série de consumo com a estimada, numa amostra de 340 horas, onde se observa uma grande proximidade.

De seguida, pretende-se calcular o resíduo da estimativa e para isso, fez-se a diferença entre a série de consumo e a sua estimativa, como se observa na equação 3.8.

$$\text{Resíduo} = \text{Consumo} - \text{Cons. Estimado} \quad (3.8)$$

Utilizando novamente o SPSS, determinaram-se os valores de ACF e PACF para a série do erro. Estes indicadores, revelam a necessidade ou não de incluir mais termos no modelo previamente estabelecido e nesse caso, seria preciso voltar a estimar os parâmetros.

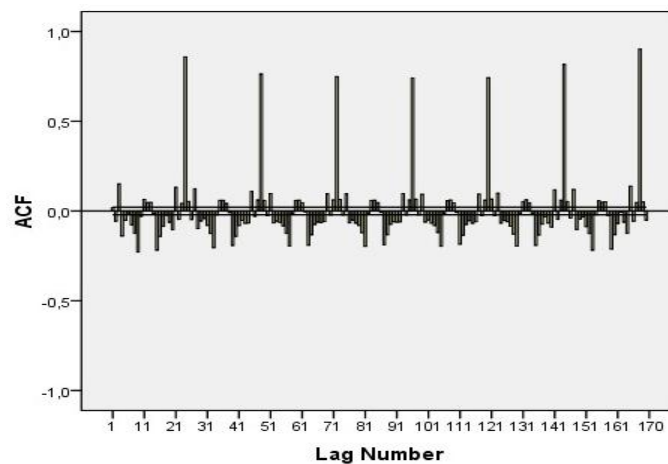


Figura 3.8 – Gráfico ACF do resíduo.

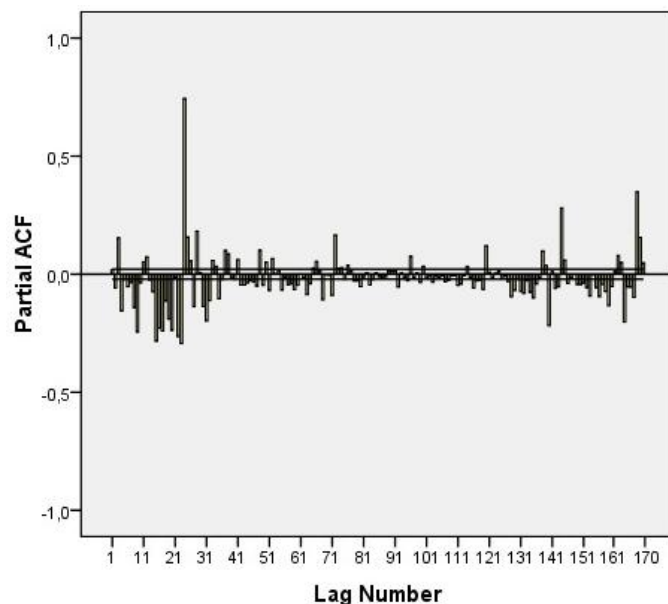


Figura 3.9 – Gráfico PACF do resíduo.

Os resíduos do modelo estimado representam estimativas de ruído branco, sendo que no caso de o modelo estar adequadamente especificado, os coeficientes de

auto-correlação dos resíduos não deverão ser significativos. A análise da figura 3.8 do gráfico ACF demonstra que os resíduos revelam um comportamento não aleatório, com destaque nos *lags* 24, 48, 72, 96, 120, 144 e 168 o que significa que a série tem sazonalidade AR de 1 dia (24 horas). Desta forma, a análise dos resíduos indica que o modelo poderá ser melhorado pois não considerava a informação relativa à sazonalidade AR da série de consumo.

Nesse caso, segundo os procedimentos da metodologia Box-Jenkins referidos anteriormente, pretende-se voltar a estimar os parâmetros para o modelo em estudo, considerando agora a informação relativa à sazonalidade.

3.5 Consideração da sazonalidade

Assim, o modelo a considerar será baseado na equação 2.5 descrita anteriormente juntamente com a informação relativa à sazonalidade de 24 horas.

$$(1-B)(1-\phi_1B-\phi_2B^2)(1-\Phi_1B^{24})X_t = a_t \quad (3.9)$$

Desenvolvendo a equação 3.9 resulta:

$$(1-\phi_1B-\phi_2B^2-B+\phi_1B^2+\phi_2B^3-\Phi_1B^{24}+\phi_1\Phi_1B^{25}+\phi_2\Phi_1B^{26}+\Phi_1B^{25}-\phi_1\Phi_1B^{26}-\phi_2\Phi_1B^{27})X_t = a_t \quad (3.10)$$

$$X_t = (1+\phi_1)*X_{t-1} + (\phi_2-\phi_1)*X_{t-2} - \phi_2X_{t-3} + \Phi_1*X_{t-24} - (\phi_1\Phi_1 + \Phi_1)*X_{t-25} + (\phi_1\Phi_1 - \phi_2\Phi_1)*X_{t-26} + \phi_2\Phi_1*X_{t-27} + a_t \quad (3.11)$$

Com a equação 3.11 definida, pretende-se estimar os parâmetros para o modelo proposto, onde se irá utilizar o método dos mínimos quadrados com o “solver” do Excel. Na estimação dos coeficientes, considera-se que $a_t = 0$.

Os valores obtidos para os coeficientes foram:

$$\phi_1 = 0,7190$$

$$\phi_2 = -0,2177$$

$$\Phi_1 = 0,8783$$

Para o somatório dos erros ao quadrado ($\sum a_t^2$) o valor determinado foi de 4,0838E+08. Deste modo, substituindo os valores dos coeficientes na equação 3.11, resulta a equação 3.12.

$$\hat{X}_t = 1,7190 \cdot X_{t-1} - 0,9367 \cdot X_{t-2} + 0,2177 \cdot X_{t-3} + 0,8783 \cdot X_{t-24} - 1,5099 \cdot X_{t-25} + 0,8227 \cdot X_{t-26} - 0,1912 \cdot X_{t-27} + a_t \quad (3.12)$$

Por observação da figura 3.10 onde se compara a série de consumo e a sua estimativa com o modelo sazonal, percebe-se que existe semelhança entre ambas tal como acontecia com o modelo anterior. A diferença é que o erro quadrático total (somatório dos erros ao quadrado) do modelo sazonal estimado foi menor, como se pode observar na Tabela 3.1.

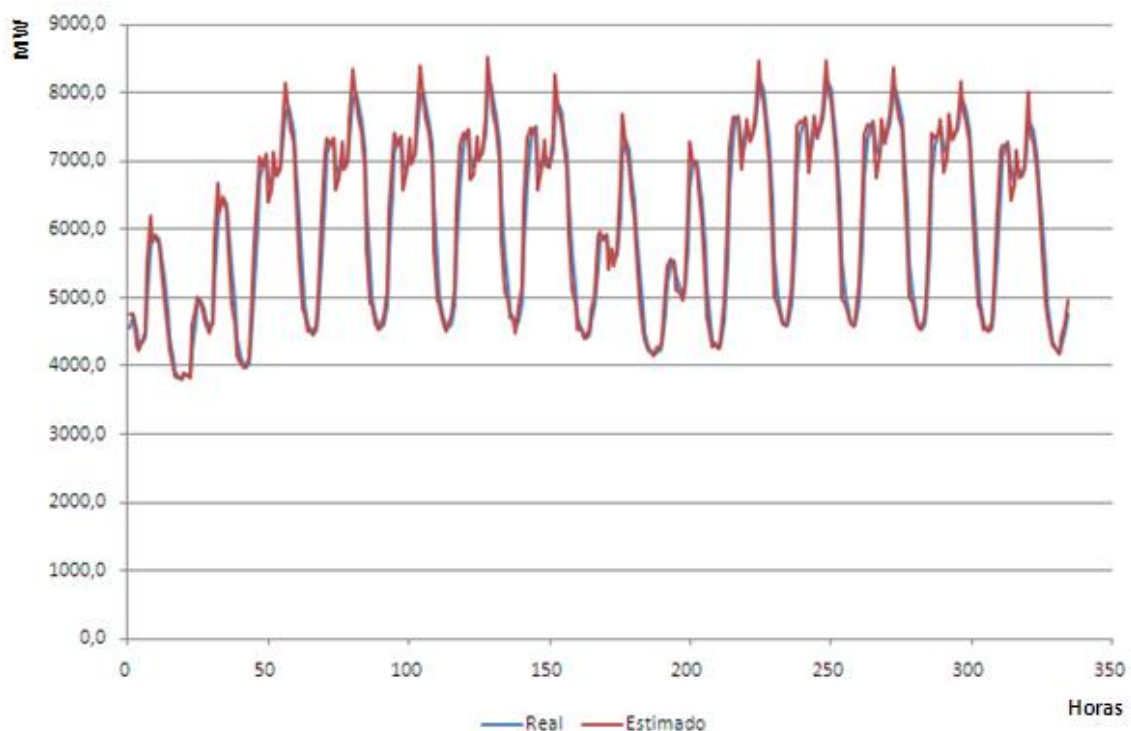


Figura 3.10 – Comparação do Consumo Real com o Estimado.

Como foi referido, o modelo sazonal estimado resultou num erro menor do que o do modelo estudado inicialmente, pelo que o próximo passo será o de avaliar os modelos estudados e realizar a análise ao seu desempenho face aos critérios definidos anteriormente, terminando assim os procedimentos da metodologia Box-Jenkins.

Deste modo, concluir-se-á sobre os valores associados aos erros de cada um dos modelos e sobre os critérios calculados para cada um, de forma a se poder obter

uma referência em relação ao desempenho e prosseguir assim, os estudos utilizando as outras técnicas nos próximos capítulos deste trabalho.

3.6 Avaliação do desempenho do modelo

Inicialmente estimou-se um modelo de previsão AR, onde se considerou que para o aperfeiçoar se incluiria a informação de sazonalidade AR. Estudaram-se 2 modelos sazonais distintos, com informação sazonal AR diferente, onde um dos modelos inclui 1 parâmetro sazonal AR enquanto o outro considera 2.

Na tabela 3.1 comparam-se os vários modelos de previsão estudados através de critérios que visam avaliar o seu desempenho, onde se retiram as respetivas conclusões.

Modelos de Previsão	Valor dos parâmetros	$\sum a_t^2$	RMSE	MAPE
AR	$\phi_1 = 0,9384$ $\phi_2 = -0,3266$	5,8674E+08	258,9	3,43%
Sazonal AR(1)	$\phi_1 = 0,7190$ $\phi_2 = -0,2177$ $\Phi_1 = 0,8783$	4,0838E+08	216,21	3,01%
Sazonal AR(2)	$\phi_1 = 0,7192$ $\phi_2 = -0,2190$ $\Phi_1 = -0,0061$ $\Phi_2 = 0,8775$	9,3087E+08	326,5	4,60%

Tabela 3.1 – Comparação dos vários modelos de previsão.

Por observação dos resultados apresentados na tabela 3.1, conclui-se que o modelo sazonal AR(1) foi o que obteve melhor desempenho, em todos os critérios calculados. Este modelo ao considerar a sazonalidade AR de 1 dia (24 horas), aperfeiçoou o modelo inicial resultando em menores valores para os erros. O outro modelo sazonal estudado considerou 2 parâmetros AR de 6 e 12 horas, no entanto estes geraram mais erros tornando-o o modelo com pior desempenho dos 3.

Comparando os 3 modelos estudados através dos vários critérios determinados, o objetivo é o de minimizar os valores dos erros e assim obter uma referência de desempenho para o melhor modelo. Deste modo, conclui-se que dos

modelos estudados o sazonal AR(1) é o que reúne melhores resultados, pelo que será considerado como referência, para os próximos estudos.

3.7 Conclusão

Neste capítulo, apresentaram-se os diversos procedimentos a efetuar com a metodologia Box-Jenkins aplicada à série de consumo considerada. Ao avaliar o desempenho dos modelos estudados, o sazonal AR(1) revelou ser o melhor, pelo que será o utilizado como referência para os estudos dos próximos capítulos.

Capítulo 4

Previsão com Redes Neurais

Neste capítulo, o objetivo é o de estruturar uma RN para um modelo de previsão da série de consumo em estudo, definindo as entradas e saídas respetivas para a RN a estudar. Posteriormente pretende-se comparar os resultados alcançados, tendo em conta os critérios definidos de avaliação ao desempenho dos modelos.

4.1 *Software* utilizado

Para o desenvolvimento deste trabalho, foi utilizado o *software Microsoft Excel* como ferramenta de filtragem e tratamento de dados e posterior análise de resultados. Para a realização da previsão foi utilizado o *software MATLAB* versão R2014a e uma das suas ferramentas/*toolboxes*, designada de *Neural Network Fitting Tool* (função “nftool”) para o treino da RN. Esta *toolbox* permite treinar uma RN utilizando para isso o algoritmo de Levenberg-Marquardt de retro propagação do erro (*Back-Propagation*).

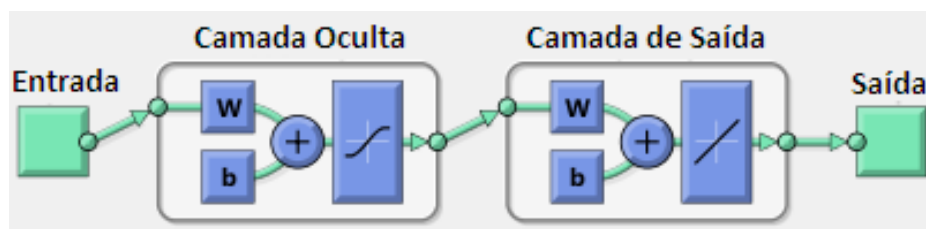


Figura 4.1 – Esquema representativo da arquitetura da RN.

Na figura 4.1 observa-se o esquema relativo à arquitetura da RN utilizada, que se configura como uma rede unidirecional (*feedforward*) de dupla camada, com neurónios ocultos ativados pela função tangente hiperbólica e neurónios de saída ativados pela função linear.

De forma a iniciar o treino da RN e depois de introduzir a série de dados na aplicação relativo às “Entradas” e “Alvos”, o conjunto de dados de treino é dividido em 3 subconjuntos, designados por “Treino”, “Validação” e “Teste” onde se pode determinar a percentagem de dados do conjunto de treino a utilizar nos outros subconjuntos. As percentagens definidas foram de 70%, 15% e 15% respetivamente.

Select Percentages

Randomly divide up the 8760 samples:

Category	Percentage	Number of Samples
Training:	70%	6132 samples
Validation:	15%	1314 samples
Testing:	15%	1314 samples

Figura 4.2 – Percentagens a definir para o treino, validação e teste da RN.

Os dados relativos ao treino vão sendo apresentados à RN e esta vai ajustando os pesos de acordo com o erro. O conjunto de validação é utilizado para monitorização da evolução do erro num conjunto independente do de treino e serve como medida de paragem do treino. Este para automaticamente, quando a generalização deixa de melhorar como indicado pelo aumento no erro médio quadrático da amostra de validação. Em relação ao conjunto de teste não tem consequência no treino, pelo que providencia uma medida independente da performance da RN depois do treino ter finalizado, ou seja é uma medida da capacidade de generalização da RN.

O erro médio quadrático é utilizado pelo MATLAB como critério de desempenho, que o programa tenta minimizar.

De forma a melhorar a performance da RN, devem ser testadas diversas arquiteturas de RN, variando o número de neurónios ocultos, para assim atingir resultados com o menor erro de generalização possível.

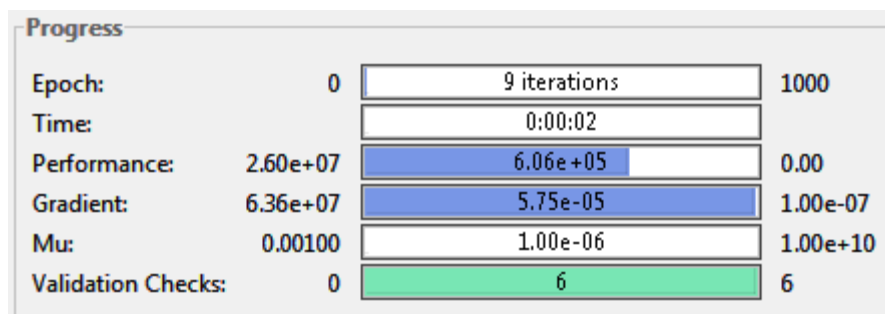


Figura 4.3 – Parâmetros de progresso durante o treino da RN.

Durante o treino da RN, é possível observar os parâmetros de progresso do algoritmo de Levenberg-Marquardt que o *software* utiliza, nomeadamente:

- Epoch, representa o número de épocas a treinar e apresenta o número de iterações;
- O tempo de treino despendido;
- A performance alcançada;
- O valor para o gradiente;
- Mu representa o *momentum* da RN;

- O número máximo de aumentos do erro no conjunto de validação: 6.

Quando o processo de treino termina, pode-se voltar a treinar a RN e assim obter resultados que possam ser um pouco melhores. A avaliação ao desempenho da rede permite apurar a sua performance e no caso de o querer melhorar, pode-se alterar o número de neurónios ocultos para um novo treino, podendo desta forma atingir melhores resultados.

4.2 Aplicação de RN em estudos de previsão

Com o auxílio do *software* MATLAB e da sua *toolbox* “nftool”, pretende-se treinar uma RN da série de dados relativa ao consumo, com o intuito de melhorar o desempenho alcançado com o modelo sazonal estudado anteriormente.

Assim, para o 1º modelo da RN a estudar, consideram-se como entradas os mesmos parâmetros utilizados pelo modelo sazonal de referência, para que seja possível a comparação de desempenho entre as duas técnicas de previsão. Na tabela 4.1, apresentam-se as variáveis consideradas para os estudos a realizar com RN.

Entradas				Alvo
Xt-1	Xt-2	Xt-24	Xt-168	Xt

Tabela 4.1 – Variáveis a considerar no treino da RN (1º modelo).

Para se determinar o número de neurónios ocultos a utilizar na RN, considera-se a sua estrutura que é composta por 2 camadas e que “n” corresponde ao número de variáveis de entrada, pelo que se obtém a equação 4.1 [8]. De notar, que para se atingir a melhor performance da RN devem-se realizar simulações com valores diferentes para os neurónios ocultos.

$$2 * n + 1 \quad (4.1)$$

Na figura 4.4 observa-se o esquema da RN definida para o treino, onde as 4 entradas resultam em 9 neurónios ocultos, segundo a equação 4.1.

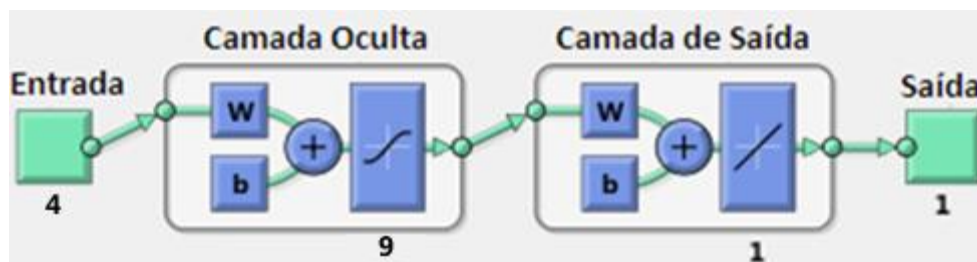


Figura 4.4 – Esquema da RN definida para o treino.

Efetuada o treino da RN é possível obter os resultados, representados na figura 4.5 para o coeficiente de correlação “R”. Este mede a correlação entre os valores de saída da RN e os valores reais, pelo que se pretende que o seu valor seja o mais próximo de 1. Pelos resultados atingidos de 0,99 conclui-se que existe uma forte correlação dos dados, facto que se observa na figura 4.5.

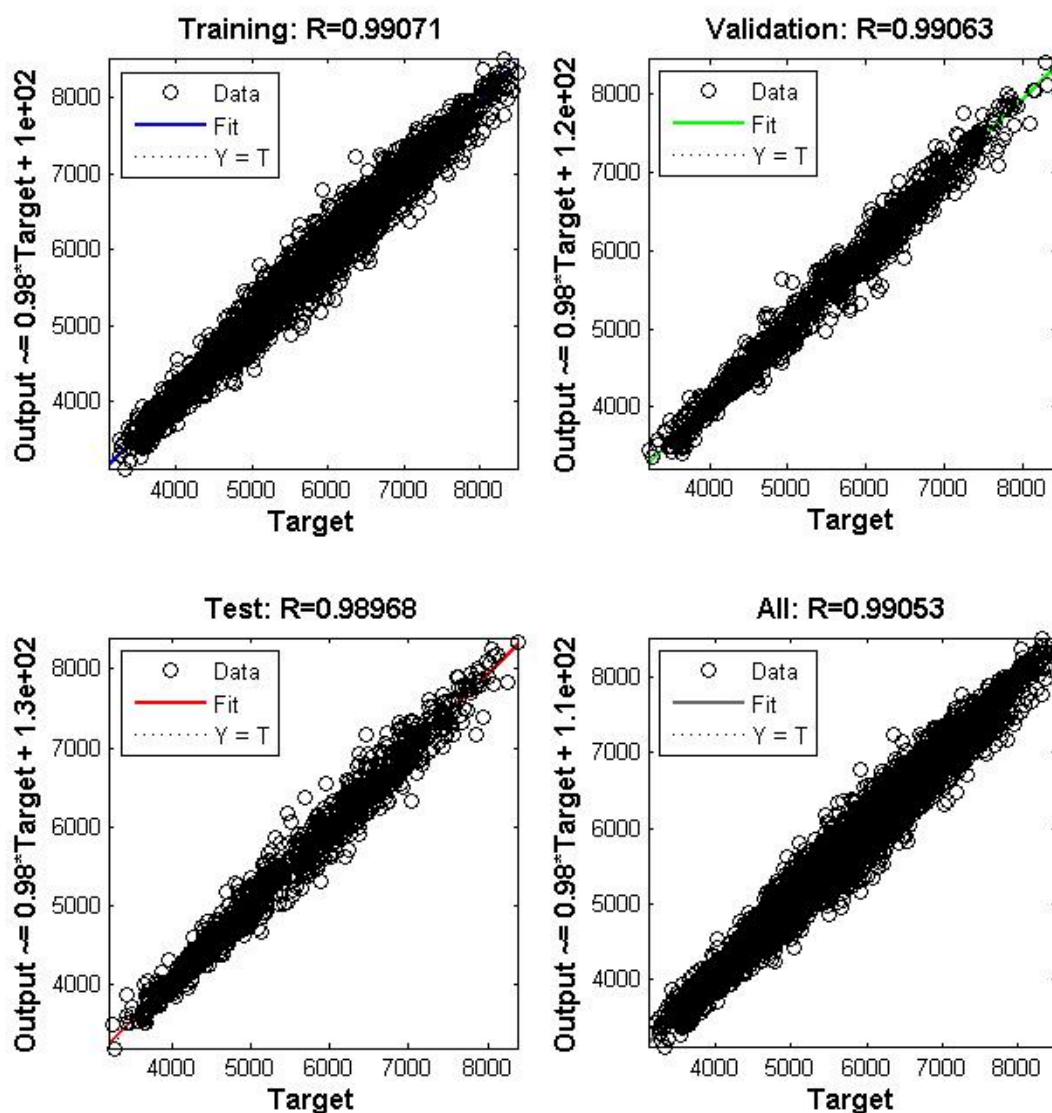


Figura 4.5 – Resultados para o coeficiente de correlação “R” do treino, validação e teste da RN.

Na figura 4.6 apresentam-se os parâmetros de progresso da RN durante o treino efetuado.

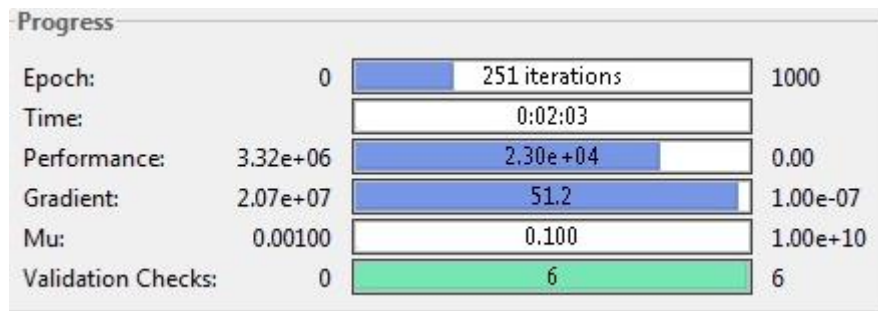


Figura 4.6 – Parâmetros de progresso durante o treino da RN.

Em seguida, pretende-se efetuar a previsão com a RN treinada. Os resultados da previsão são fornecidos pelas saídas obtidas pelo treino da RN.

De forma a avaliar a qualidade da previsão efetuada pela RN, na figura 4.7 compara-se a série de dados do consumo com os resultados obtidos pela RN, numa amostra de 340 horas.

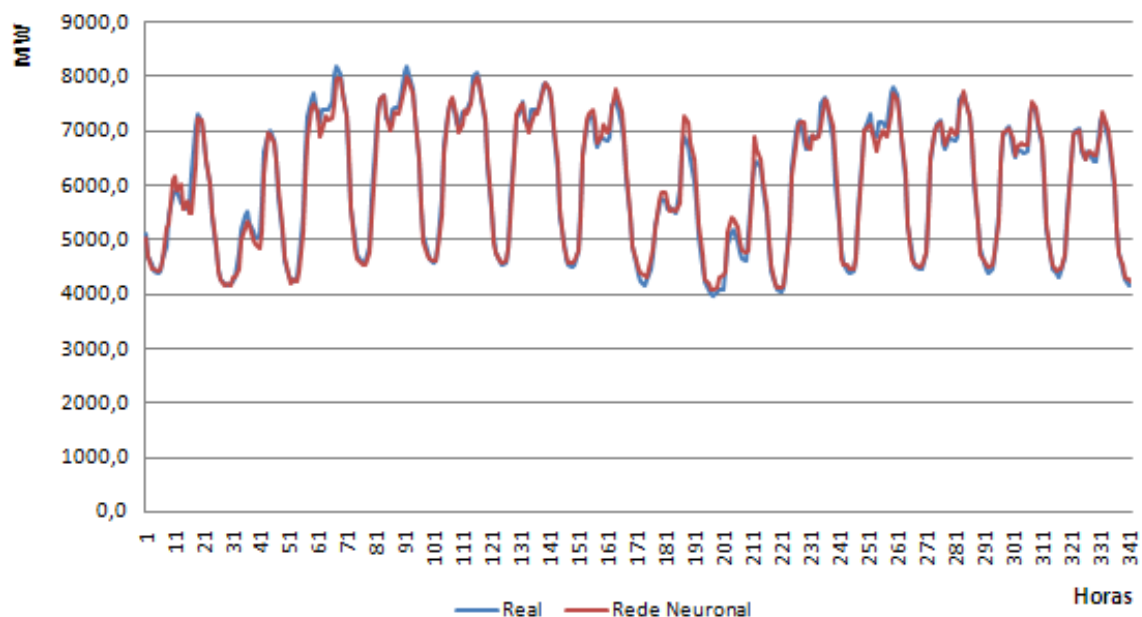


Figura 4.7 – Comparação do Consumo Real com a Rede Neuronal.

Quanto ao somatório dos erros ao quadrado ($\sum a_t^2$), um dos critérios de avaliação ao desempenho dos modelos, o resultado obtido foi de 2,2428E+08. No final deste capítulo, apresentam-se todos os resultados que permitem aferir sobre o desempenho do método, como é o caso do RMSE e MAPE.

4.3 Estandardização dos dados de entrada

O conceito de estandardização dos dados define-se como um processo de ajuste de escalas, que melhora assim a eficiência de treino da RN. Consiste numa equalização das escalas das variáveis de entrada e saída para gamas no intervalo entre -1 e 1, conseguindo desta forma um melhor ajuste da gama das variáveis às funções de ativação, equalizando também as escalas entre as diversas variáveis [7].

Um dos métodos mais conhecidos designa-se por “Min-Máx” e afigura-se como o melhor quando se conhece o mínimo e o máximo da variável em estudo, como é o caso da série de consumo deste trabalho.

Designa-se por destandardização o processo inverso, que transforma a escala entre -1 e 1 das variáveis estandardizadas na escala real.

Supondo o caso em que \min_a e \max_a são o mínimo e o máximo do conjunto de entrada “a”, a estandardização Min-Máx mapeia o valor “v” de “a” na escala [\min_a , \max_a] em “v” de “A” na escala [\min_A , \max_A]. Nas equações 4.2 e 4.3 em seguida, apresentam-se as fórmulas de cálculo para v' e v :

Estandardização:

$$v' = \frac{v - \min_a}{\max_a - \min_a} * (\max_A - \min_A) + \min_A \quad (4.2)$$

Destandardização:

$$v = \frac{v' - \min_A}{\max_A - \min_A} * (\max_a - \min_a) + \min_a \quad (4.3)$$

Para os valores de \min_A e de \max_A é habitual assumir os valores de -1 e 1, respetivamente.

Similarmente ao que foi feito para iniciar o treino da RN anterior, consideram-se as mesmas variáveis para se proceder a um novo treino de RN, tal como se observa na tabela 4.1. Desta forma, procedeu-se à estandardização das variáveis de entrada tendo em conta a equação 4.2, pretendendo-se treinar a RN para posterior destandardização dos resultados atingidos pela equação 4.3 e assim, avaliar o desempenho do método.

Como anteriormente, após o treino da RN observam-se os resultados para o coeficiente de correlação “R”, para os parâmetros de progresso e compara-se a série de dados de consumo com o obtido pela RN numa amostra de 340 horas, nas figuras 4.8, 4.9 e 4.10 respetivamente.

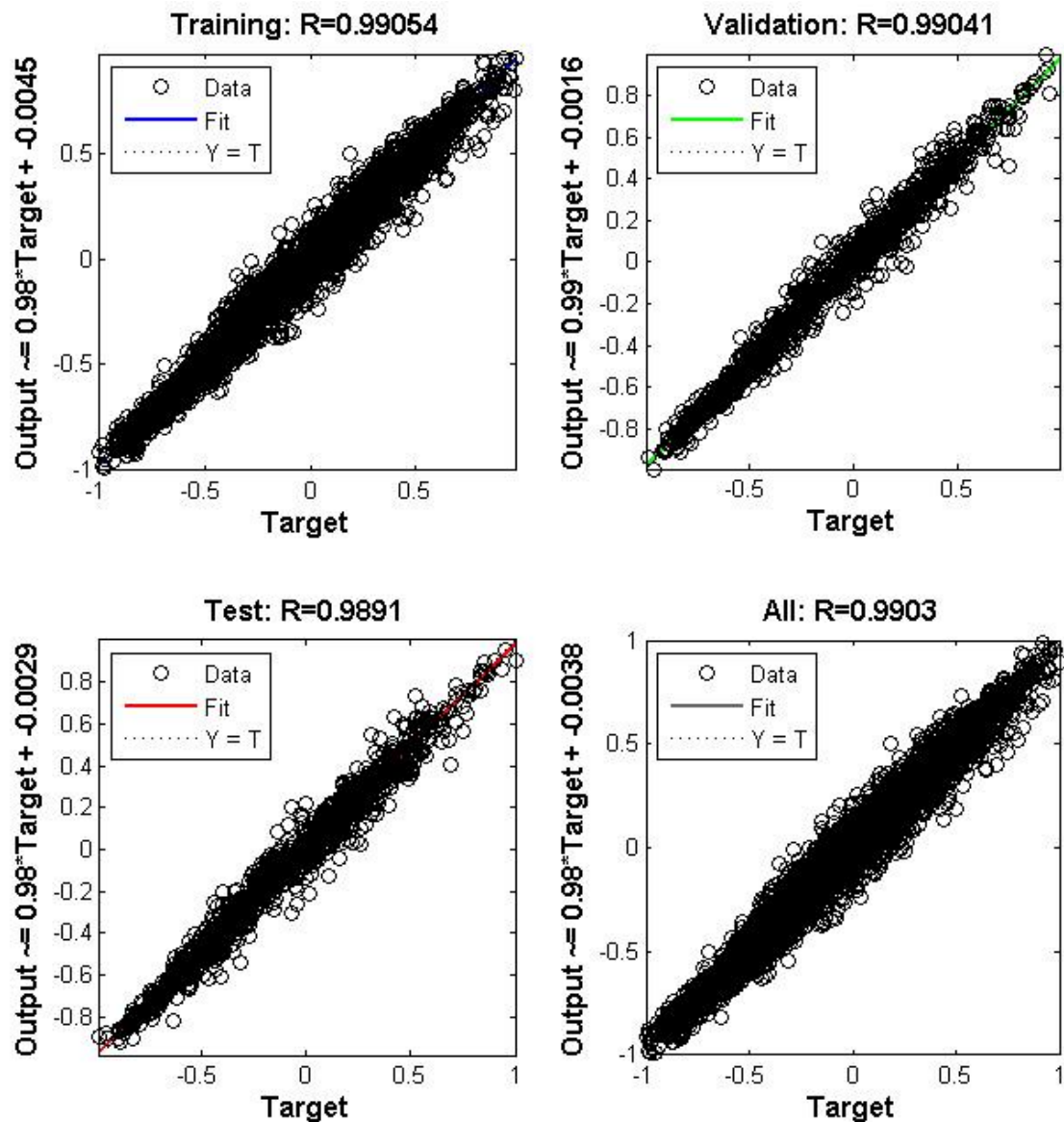


Figura 4.8 – Resultados para o coeficiente de correlação “R” do treino, validação e teste da RN estandardizada.

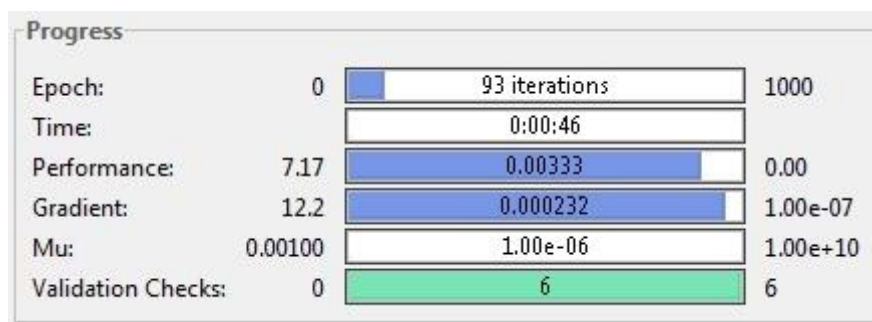


Figura 4.9 – Parâmetros de progresso durante o treino da RN estandardizada.

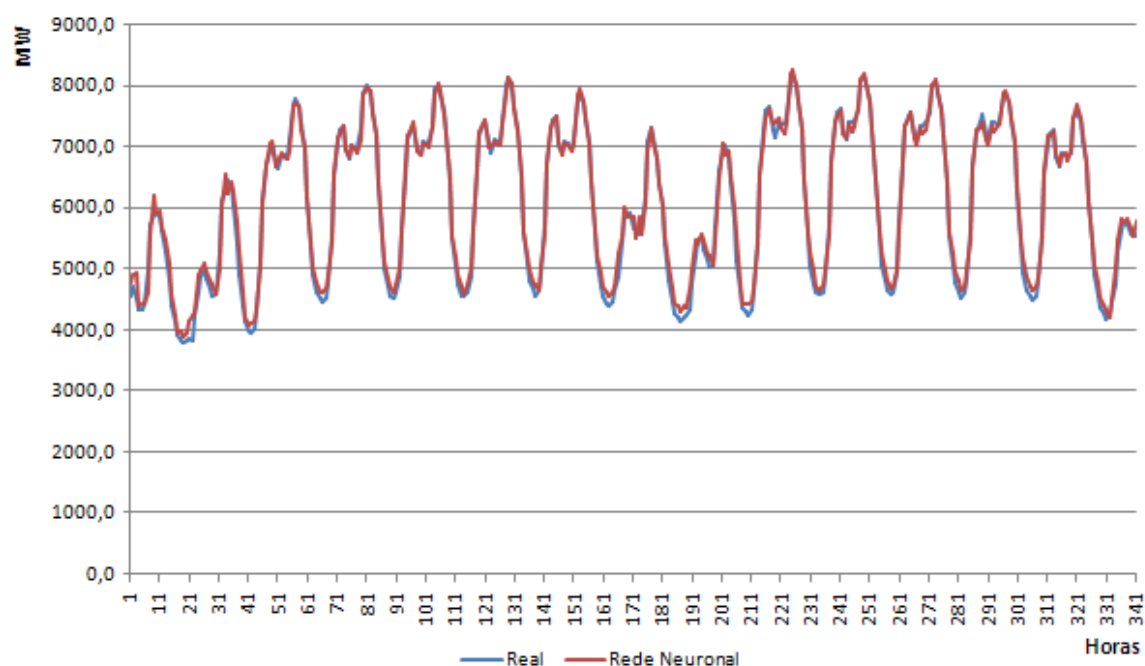


Figura 4.10 – Comparação do Consumo Real com a Rede Neuronal Estandarizada.

Quanto ao somatório dos erros ao quadrado ($\sum a_t^2$), um dos critérios de avaliação ao desempenho dos dados estimados, o resultado atingido foi de 2,1679E+08.

De forma semelhante ao estudo efetuado para o 1º modelo da RN, considera-se um 2º modelo constituído por outras variáveis de entrada, como se observa na tabela 4.2. Pretende-se assim considerar uma nova entrada, que corresponde a 336 horas (número de horas em 2 semanas) e que é pertinente acrescentar ao estudo, devido à sazonalidade da série de consumo.

Entradas					Alvo
Xt-1	Xt-2	Xt-24	Xt-168	Xt-336	Xt

Tabela 4.2 – Variáveis a considerar no treino da RN (2º modelo).

Os vários resultados que permitem aferir sobre o desempenho dos modelos, apresentam-se na tabela 4.3.

4.4 Caracterização do consumo face ao dia de semana e hora

Em estudos de previsão do consumo de energia elétrica utilizando as RN e face às características relevantes que a variação do consumo apresenta, nomeadamente em relação às horas do dia e ao dia de semana (por exemplo) onde se observam padrões na sua periodicidade, pretende-se estudar a sua influência nos modelos de previsão de RN utilizados. Desta forma, as variáveis relativas às horas e ao dia de semana podem ser incluídas como auxiliares nas entradas das RN.

Considerando o dia da semana, a razão principal para se considerar esta variável reside no facto de se verificar um consumo de energia elétrica diferenciado, entre um dia de semana “útil” e um dia de fim-de-semana, algo que facilmente se observa nas figuras 4.7 e 4.10.

Esta diferença no consumo pode ser explicada devido a grande parte da indústria e algum comércio, principalmente ao domingo, estarem encerrados resultando num menor consumo em dias de fim-de-semana.

Em relação à hora, facilmente se constata através da observação da sua evolução no período de um dia (figuras 1.1 e 1.2), que existe um padrão que relaciona os diferentes períodos horários com o consumo. Nota-se que nas primeiras horas do dia, o consumo permanece mais baixo, até que começa a subir logo pelo início da manhã continuando até pouco depois do meio da tarde, onde decai ligeiramente para de seguida subir a partir do final da tarde, atingindo o pico por volta da hora de jantar. Em dias “úteis” de segunda a sexta, este processo é cíclico evidenciando-se o seu padrão.

A utilização de funções trigonométricas afiguram-se assim, como o ideal para a definição de variáveis cíclicas, usando as entradas baseadas nas funções seno e cosseno. O dia de semana ou a hora a prever, são interpretadas de uma forma periódica e a utilização das funções trigonométricas justificam-se por permitirem especificar cada dia ou hora de forma distinta. Assim, apresentam-se as equações 4.4 e 4.5 para o dia de semana (DS) e hora, respetivamente [7].

$$DS \text{ sen} = \text{sen}\left(\frac{2\pi d}{7}\right) \quad DS \text{ cos} = \text{cos}\left(\frac{2\pi d}{7}\right) \quad (4.4)$$

$$Hora \text{ sen} = \text{sen}\left(\frac{2\pi h}{24}\right) \quad Hora \text{ cos} = \text{cos}\left(\frac{2\pi h}{24}\right) \quad (4.5)$$

4.4.1 Consideração dos feriados

Do ponto de vista de um estudo de previsão, os feriados são um fenómeno que afetam o consumo de energia elétrica e que se afastam dos parâmetros normais estabelecidos pela periodicidade temporal. Principalmente porque existem feriados que ocorrem sempre no mesmo dia do ano, os que mudam de dia em anos diferentes (por exemplo devido às fases da lua, como o Carnaval e a Páscoa) e os feriados regionais que dependem da localidade (por exemplo o Santo António e o São João). Além disso, os feriados ainda possuem outra particularidade, que é o facto de que nem todos se comportam da mesma maneira no que diz respeito ao consumo de energia elétrica. Feriados como o 1 de janeiro e o Carnaval possuem diagramas de carga completamente diferentes de outros feriados como o Natal ou a Páscoa, onde no primeiro caso existe um aumento significativo do consumo de energia elétrica de madrugada, devido às celebrações que normalmente ocorrem nesses dias.

Os feriados são por definição, dias oficiais em que não se trabalha devido à comemoração de uma determinada ocasião. No entanto, o conceito de não se trabalhar nos feriados é algo vago, visto que em alguns setores da sociedade se trabalha nesses dias, como o comércio e alguma indústria. Os feriados são um fator que perturba a periodicidade do consumo semanal, até porque quando se aproximam dos fins-de-semana por vezes originam pontes o que também afeta o padrão de consumo.

Por isso é que devido à incerteza que está associada aos feriados e mesmo em relação ao dia que se segue, quando se realizam estudos de previsão a sua influência provoca efeitos negativos na performance do modelo. Esses efeitos ocorrem tanto no dia do feriado, como também nos dias seguintes que dependem dos seus valores. Normalmente o que acontece, é que as previsões em dias de feriado têm tendência a erros por excesso pois o consumo é menor do que o que se espera. Por outro lado, as previsões dos dias que se seguem, que utilizam os valores do dia de feriado têm a tendência para erros por defeito, ou seja os valores previstos são menores do que os reais.

Como exemplo, observa-se na figura 4.11 o consumo do feriado de dia 27 de março de 2005, onde se compara o real e o previsto pelo 2º modelo de RN estandardizado. Desta forma, pretende-se demonstrar a tendência que normalmente ocorre nos feriados onde o consumo previsto excedeu o real, o que resultou num dos dias com maior erro de previsão do ano.

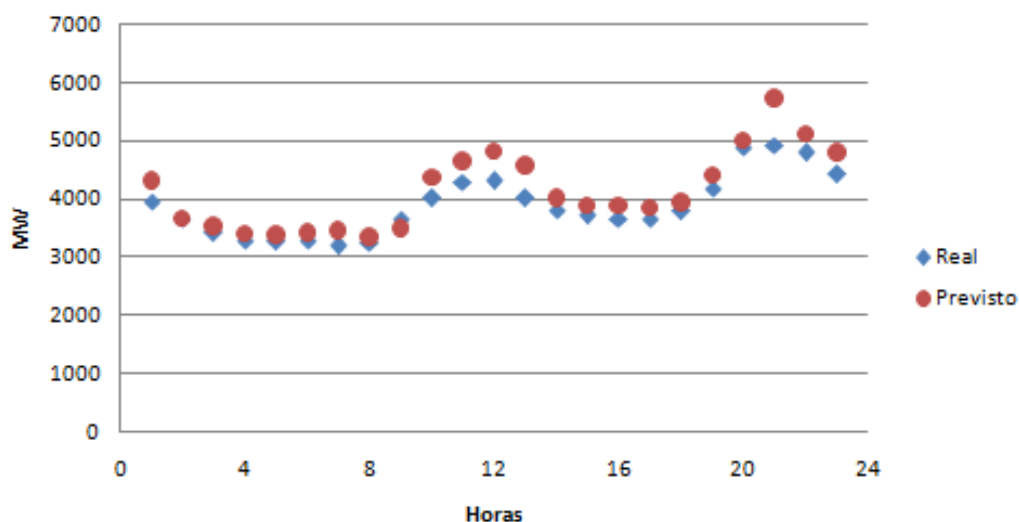


Figura 4.11 – Comparação do consumo real e previsto no feriado de 27 de março 2005.

Desta forma, torna-se complexo estabelecer uma RN que consiga detectar todas estas variações e que seja capaz de prever as diferentes alterações. Uma forma de ultrapassar este obstáculo, é o de considerar que o consumo de um dia feriado, pode ser comparado a um dia de fim-de-semana (sábado ou domingo), visto que os seus diagramas de carga aproximam-se dos padrões uns dos outros.

Face ao exposto, considera-se um 3º modelo de RN a estudar em 3 situações distintas:

- a RN considera as entradas do modelo anterior (o segundo), acrescentando as variáveis auxiliares dia de semana (DS) e hora;
- neste estudo, considera-se o mesmo da etapa anterior com a diferença que os feriados são tratados como se fossem sábados (“F-S”);
- semelhante à etapa anterior, com a diferença que os feriados são tratados como se fossem domingos (“F-D”).

Os resultados destes estudos apresentam-se na tabela 4.3, onde se conclui sobre o desempenho dos métodos.

4.5 Avaliação do desempenho do modelo

Neste capítulo estudaram-se 3 modelos de RN, onde o primeiro considerava as variáveis do modelo sazonal AR(1) referenciado anteriormente, pretendendo-se comparar os resultados atingidos em igualdade de circunstâncias, enquanto o segundo acrescentou outra entrada na RN a estudar, de forma a poder melhorar o desempenho do modelo. Finalmente o 3º modelo, considerou as variáveis cíclicas do dia de semana

e hora numa primeira fase, enquanto que as variantes ao 3º modelo designadas por “F-S” e “F-D” além de as incluírem, referem-se à situação dos dias feriados terem sido tratados como se fossem sábados e domingos, respetivamente.

Na tabela 4.3 comparam-se os vários métodos de previsão estudados através dos critérios que visam avaliar o seu desempenho.

Método de Previsão	$\sum a_t^2$	RMSE	MAPE
Sazonal AR(1)	4,0838E+08	216,21	3,01%
RN (1º modelo)	2,2428E+08	161,56	2,19%
RN Estand. (1º modelo)	2,1679E+08	158,84	2,13%
RN (2º modelo)	2,0925E+08	157,61	2,12%
RN Estand. (2º modelo)	1,8999E+08	150,18	2,02%
RN (3º modelo)	1,3885E+08	128,39	1,68%
RN Estand. (3º modelo)	1,2685E+08	122,71	1,61%
RN (3º modelo F-S)	1,3872E+08	128,32	1,67%
RN Estand. (3º modelo F-S)	1,1803E+08	118,37	1,59%
RN (3º modelo F-D)	1,1732E+08	118,02	1,58%
RN Estand. (3º modelo F-D)	1,1099E+08	114,79	1,53%

Tabela 4.3 – Comparação de desempenho.

Por observação dos resultados apresentados na tabela 4.3, observa-se que face aos diferentes modelos utilizados os vários resultados foram melhorando. A diferença mais notória resulta da mudança de técnica de previsão, onde a alteração do método de Box-Jenkins para as RN (1º modelo) foi significativo e demonstrativo do desempenho da RN. Com o 2º modelo de RN, acrescentou-se a entrada relativa às 336

horas (2 semanas) que se relevou importante para o desempenho da RN. Além disso, a standardização dos dados demonstrou a sua validade pois obteve os melhores resultados nos diferentes modelos.

A segunda maior diferença nos resultados, prende-se com a utilização das variáveis auxiliares dia de semana e hora, que devido à sua característica cíclica melhorou o desempenho de forma relevante. Ao considerar os feriados nas duas variantes estabelecidas, demonstrou-se que a situação de os comparar aos domingos, resultou na diminuição do erro o que valida este método.

4.6 Conclusão

Neste capítulo, apresentaram-se os diversos procedimentos a efetuar com as RN aplicadas à série de consumo considerada. Realizaram-se diversos estudos distintos, destacando-se os resultados dos modelos standardizados e os que consideravam as variáveis auxiliares DS e hora e também os feriados.

Capítulo 5

Previsão com SVM

Neste capítulo, o objetivo é o de estruturar um modelo de previsão da série de consumo em estudo, utilizando para isso as SVM. Posteriormente pretende-se comparar os resultados atingidos com o objetivo de determinar o modelo com melhor desempenho, tendo em conta os critérios de avaliação definidos anteriormente.

5.1 *Software* utilizado

Analogamente às RN, no desenvolvimento deste trabalho foi utilizado o *software Microsoft Excel* como ferramenta de filtragem e tratamento de dados, enquanto que para se proceder ao treino das SVM e à previsão foi utilizado o *software* MATLAB versão R2014a, através da *toolbox* desenvolvida por Steve Gunn.

Com a utilização destas ferramentas, pretende-se treinar e testar diferentes modelos SVM tendo em consideração as diferentes combinações dos seus parâmetros e os consequentes efeitos na eficácia do modelo de previsão. O objetivo é o de determinar a combinação de parâmetros que resulte no modelo com melhor desempenho.

Tal como qualquer modelo de aprendizagem supervisionada, numa fase inicial é necessário treinar a SVM, para de seguida se efetuar a previsão. No que diz respeito ao treino da SVM, este é feito por utilização da função “*svr*” e dos seus parâmetros: *X*, *Y*, função kernel, *C*, função penalização e ϵ .

O parâmetro “*X*” refere-se à matriz de dados de treino que corresponde às entradas a considerar no estudo e “*Y*” é a matriz de dados referente ao alvo do treino.

De forma a otimizar o processo de treino, existem outros parâmetros que se podem acrescentar à simulação, como é o caso das funções kernel. Normalmente, o espaço dimensional original das entradas não é linearmente relacionável com as variáveis a prever e nestes casos é utilizada uma função kernel que transforma num espaço de maior dimensão, de modo a ser possível contornar as não linearidades do espaço dimensional inicial. Assim, podem ser utilizadas diferentes funções kernel no processo de treino, sendo as mais comuns as funções lineares, polinomiais e de base radial. Neste estudo será utilizada a função de base radial pois além de ser a mais utilizada/indicada em estudos de previsão de carga com SVM, foi a que obteve melhores resultados face às diferentes funções kernel testadas.

Além disso, uma SVM é definida a partir da constante “C” que está relacionada com a função objetivo a ser minimizada ao longo do treino e da sua função de penalização com tolerância ϵ (como se viu na figura 2.7).

Caso existam valores em falta nas matrizes “X” e “Y” durante o processo de treino, a função reconhece esses valores em falta e ignora as linhas correspondentes da outra matriz.

Em relação aos parâmetros de saída resultado do treino da SVM, são eles:

- vetores de suporte, referem-se a uma matriz de dados onde cada linha corresponde a um vetor de suporte no espaço de dados normalizado;
- alfa, é o vetor de pesos para os vetores de suporte;
- desvio (também designado por *bias*), corresponde à interceção do hiperplano que separa os dois grupos no espaço de dados normalizado.

Para se proceder à fase de previsão, utiliza-se a função “svroutput” juntamente com a SVM treinada utilizando os parâmetros: X, testeX, função kernel, alfa e desvio.

O parâmetro “X” refere-se à matriz de dados de treino e “testeX” é a matriz de dados de teste, ambas as matrizes são da mesma dimensão. Quanto aos outros parâmetros, são provenientes da fase de treino e são usados para a previsão.

Como resultado, obtém-se as saídas respeitantes à previsão efetuada com os diferentes parâmetros definidos no modelo SVM configurado. Desta forma pretende-se apurar o modelo SVM com a configuração de parâmetros que resulte na previsão com menor erro.

5.2 Resultados

Com os diferentes modelos SVM configurados, o objetivo é o de comparar o seu desempenho com os das RN estudados. Para o desenvolvimento deste estudo, consideram-se as mesmas entradas do 1º modelo de RN (ver tabela 4.1), com os dados estandardizados.

Desta forma, na tabela 5.1 observam-se os diferentes modelos SVM estudados, os parâmetros utilizados e o MAPE de cada.

Método de Previsão	C	ϵ	MAPE
SVM (1º modelo)	2,0	1,0	4,01%
SVM (2º modelo)	0,1	0,5	2,71%
SVM (3º modelo)	0,01	0,1	2,09%

Tabela 5.1 – Comparação dos diferentes modelos SVM.

Por observação da figura 5.1, percebe-se que existe uma grande proximidade entre o consumo real e o previsto com o 3º modelo de SVM contrastando com a diferença para o 1º modelo SVM, em que se nota um erro maior.

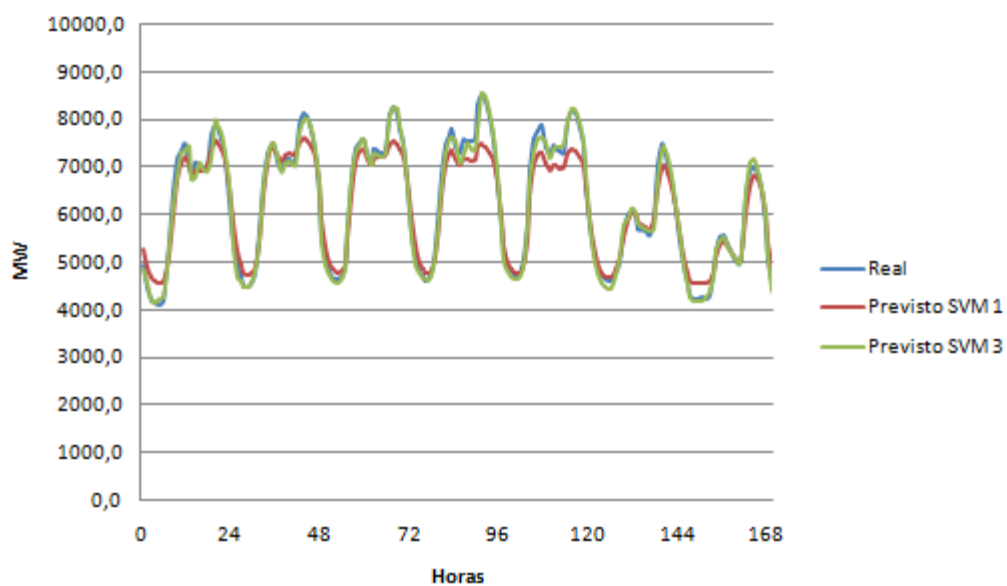


Figura 5.1 – Comparação entre o consumo real e o previsto.

Com os modelos SVM estudados e tendo se determinado o melhor, apresenta-se na tabela 5.2 a comparação dos resultados obtidos pelos diferentes métodos de previsão.

Método de Previsão	Σa_t^2	RMSE	MAPE
Sazonal AR(1)	4,0838E+08	216,21	3,01%
RN (1º modelo)	2,2428E+08	161,56	2,19%
RN Estand. (1º modelo)	2,1679E+08	158,84	2,13%
SVM (3º modelo)	1,9894E+08	154,79	2,09%

Tabela 5.2 – Comparação de desempenho.

Tendo em conta os resultados obtidos e ao se comparar o desempenho das SVM com o modelo sazonal AR(1), observa-se que as SVM obtiveram um desempenho superior. Comparando as SVM com o 1º modelo de RN estudado, verifica-se que as SVM resultaram num melhor desempenho, o que valida esta técnica como sendo a melhor escolha a utilizar em estudos de previsão.

5.3 Conclusão

Neste capítulo, apresentaram-se os procedimentos a efetuar com as SVM aplicadas à série de consumo em estudo. Ao avaliar o desempenho das SVM estudadas, observou-se que este método foi superior ao 1º modelo das RN. De futuro, seria interessante utilizar as SVM estudando o seu desempenho com a influência dos feriados, tal como foi feito com o 3º modelo de RN.

Capítulo 6

Comparação das Técnicas

Neste capítulo, pretende-se comparar os resultados das várias técnicas e determinar o desempenho do melhor caso de cada uma, de acordo com os respetivos critérios definidos nos capítulos anteriores. Além disso, é feita uma análise das vantagens e desvantagens de cada técnica.

Assim, observa-se a tabela 6.1 onde se compara o desempenho dos melhores modelos de previsão. Facilmente se percebe que o método Box-Jenkins que considerou a sazonalidade com 1 parâmetro AR, foi o que obteve pior desempenho, enquanto que o 3º modelo de RN estandardizado e que considerava os feriados como se fossem domingos em relação ao consumo, revelou ser o melhor.

Em relação ao desempenho das SVM, refere-se que com as mesmas considerações do 1º modelo de RN estandardizado o desempenho das SVM foi superior, pelo que seria pertinente estudar a influência dos feriados da mesma maneira do que foi feito com as RN (3º modelo) e assim comparar os respetivos resultados.

Método de Previsão	$\sum a_t^2$	RMSE	MAPE
Sazonal AR(1)	4,0838E+08	216,21	3,01%
SVM (3º modelo)	1,9894E+08	154,79	2,09%
RN (3º modelo F-D)	1,1732E+08	118,02	1,58%
RN Estand. (3º modelo F-D)	1,1099E+08	114,79	1,53%

Tabela 6.1 – Comparação de desempenho dos melhores modelos de cada método.

Em relação aos procedimentos de cada técnica e à facilidade ou complexidade de as trabalhar, sem dúvida que as RN demonstraram ser as mais intuitivas de utilizar. Para isso, facilitou o facto de existir uma *toolbox* específica para o Matlab (nftool), que com a ajuda de uma *interface* “amiga do utilizador”, permitiu uma compreensão

rápida do seu funcionamento e também de atingir os resultados. Relativamente aos tempos de treino utilizando as RN, verificou-se que mesmo com milhares de dados a serem processados, o treino convergia em poucos minutos (à volta de 2 minutos).

Quanto ao modelo Box-Jenkins, exigiu a execução de vários procedimentos distintos, o que implicava avanços mais demorados no desenvolvimento dos estudos efetuados.

Finalmente o modelo SVM, revelou ser um pouco menos intuitivo de utilizar comparando com as RN, pois não tinha uma *interface* gráfica que ajudasse ao desenvolvimento do trabalho, sendo necessário escrever em código Matlab qualquer utilização de uma função e qualquer opção de parametrização, que pudesse otimizar o estudo. Ou seja, apesar do potencial deste método de previsão e de depois de vários estudos/tentativas ter sido o que obteve melhores resultados (nas mesmas condições que o 1º modelo de RN), percebe-se que face às suas opções de otimização e às aplicações em que se pode utilizar, é um método muito interessante de se utilizar para futuros estudos.

Relativamente aos tempos de treino com as SVM, um processo envolvendo milhares de dados implicava um esforço computacional exigente, o que resultava em tempos de treino bastante demorados (pelo menos de trinta minutos).

Capítulo 7

Conclusão

Este trabalho pretende demonstrar os procedimentos relativos às técnicas de previsão aplicadas à série de consumo em estudo e aferir sobre o seu desempenho. Com os dados de consumo disponibilizados, o foco deste trabalho foi o de criar modelos de previsão com cada uma das técnicas em estudo (Box-Jenkins, RN e SVM), tendo em conta que o horizonte de previsão seria a hora seguinte.

Inicialmente, estudou-se o modelo Box-Jenkins com 2 parâmetros AR onde não se considerava a sazonalidade. Posteriormente, determinou-se a relevância de incluir no modelo o termo sazonal respeitante a 24 horas, o que revelou melhores resultados face ao modelo anterior.

Para o treino das RN, foram testados vários conjuntos de entrada e diversas arquiteturas variando o número de neurónios ocultos, o que permitiu otimizar o desempenho da RN.

Neste estudo confirmou-se a importância da standardização das entradas e alvos e posterior destandardização das saídas, de forma a promover a equalização das escalas das variáveis, o que contribuiu para uma melhoria nos resultados obtidos. Outra contribuição positiva, veio da utilização de variáveis trigonométricas como entradas auxiliares, para representação de variáveis cíclicas. Isto foi aplicado no 3º modelo de RN estudado, que considerava o dia de semana (DS) e a hora, que se sabe terem influência no consumo de energia elétrica devido aos padrões existentes.

Relativamente ao tratamento dos feriados, consideraram-se 2 situações distintas onde no primeiro estudo foram tratados como se fossem sábados ("F-S"), enquanto no outro estudo os feriados foram tratados como domingos ("F-D"). Os resultados demonstraram a viabilidade deste método e dos seus pressupostos, tendo havido uma melhoria face aos resultados dos modelos anteriores, com destaque para o último estudo ("F-D"). Assim, as RN demonstraram ser uma técnica muito robusta, de utilização intuitiva e com bom desempenho.

De referir que as SVM, apresentaram melhor desempenho em relação ao 1º modelo de RN o que demonstra o seu potencial em estudos de previsão, pelo que é de considerar em trabalhos futuros.

A comparação entre as técnicas usadas mostrou que o melhor desempenho foi obtido com o 3º modelo de RN standardizado, que tratava os feriados como domingos. No entanto, em igualdade de circunstâncias as SVM obtiveram melhor desempenho do que o 1º modelo de RN, o que demonstra a sua capacidade em estudos de previsão.

Trabalhos futuros

Quanto ao desenvolvimento de trabalho futuro, existirá margem para trabalhar outros pontos de interesse no contexto do tema desta tese. Por exemplo, a inclusão de dados meteorológicos (como a temperatura) como entrada nas RN, ou então considerar a variável respeitante aos meses do ano, realizar estudos de previsão com diferentes períodos temporais e estudar séries de dados de consumo mais atuais.

Em relação às SVM, faria sentido estudar com maior detalhe acerca da influência dos feriados na previsão e comparar com os resultados obtidos deste trabalho.

Seria também interessante, realizar estudos de previsão tendo o conhecimento do consumo por classe, considerando os casos do setor elétrico doméstico, comercial e industrial.

Os estudos desenvolvidos neste trabalho incidiram sobre um único conjunto de dados. Seria interessante testar as mesmas técnicas noutras séries de modo a confirmar se as conclusões obtidas nesta dissertação são generalizáveis ou se, pelo contrário, dependem dos dados usados.

Referências

- [1] Centro de Informação da REN. Disponível em <http://www.centrodeinformacao.ren.pt>. Visitado em outubro de 2013.
- [2] Relatório de dados técnicos da REN de 2012. Disponível em <http://www.centrodeinformacao.ren.pt>. Visitado em outubro de 2013.
- [3] Fidalgo, J.N., Monteiro, C. – *Notas técnicas de Séries Temporais*. Conteúdos da disciplina Técnicas de Previsão, do curso de Eletrotecnia da FEUP.
- [4] Peter J. Brockwell Richard A. Davis – *Introduction to Time Series and Forecasting*, 2ª edição, Springer.
- [5] Spyros Makridakis, Steven C. Wheelwright – *Forecasting Methods and Applications*, 3ª edição 1998, John Wiley & Sons, Nova Iorque.
- [6] Constituição de um neurónio biológico. Disponível em www.neuralpower.com. Visitado em outubro de 2013.
- [7] Fidalgo, J.N., Monteiro, C. – *Notas técnicas de Redes Neurais*. Conteúdos da disciplina Técnicas de Previsão, do curso de Eletrotecnia da FEUP.
- [8] Rojas, R. – *Neural Networks: A Systematic Introduction*, 1996, Springer-Verlag, New York Incorporated.
- [9] Demuth, H.B., Hagan, M., Beale, M. – *Neural Network Toolbox User's Guide*, edição R2013b, The MathWorks Incorporated, 2013.
- [10] Hamel, L. – *Knowledge Discovery with Support Vector Machines*, John Wiley & Sons Inc., 2009.
- [11] Douglas C. Montgomery, George C. Runger - *Estatística aplicada e probabilidade para engenheiros*, 2ª edição 2003 LTC, Rio de Janeiro.
- [12] Bo-Juen Chen, Ming-Wei Chang, Chih-Jen Lin – *Load Forecasting Using Support Vector Machines: A Study on EUNITE Competition 2001*. Departamento de Engenharia Informática, Universidade Nacional de Taiwan.
- [13] Liane Werner, José Ribeiro – *Previsão de Demanda: Uma Aplicação dos Modelos Box-Jenkins na Área de Assistência Técnica de Computadores Pessoais*, Universidade Federal do Rio Grande do Sul, abril 2003 V.10 nº 1.
- [14] Mohandes M. (2002) – *Support Vector Machines for short-term electrical load forecasting*, Elsevier - *International Journal of Energy Research*, 26 (4), p. 335-345, doi: 10.1002/er.787.

[15] Vitor Hugo Ferreira, *“Técnicas de Regularização de Modelos Neurais Aplicadas à Previsão de Carga a Curto Prazo”*, fevereiro de 2005, Rio de Janeiro.